1. Федеральное государственное автономное образовательное учреждение высшего образования
2. **«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**
3. **«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**
4. **Факультет экономических наук**
5. **Образовательная программа «Экономика»**

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

Моделирование кредитного рейтинга российских эмитентов ВДО

Выполнил:

Студент группы БЭК199

Поспелов Артем Алексеевич

Научный руководитель:

к. э. н., доцент

Столяров Андрей Иванович

|  |  |
| --- | --- |
|  |  |

Москва, 2023

**Оглавление**

[Введение 4](#_Toc134490416)

[Глава I. Теоретическая база исследования и обзор рынка 8](#_Toc134490417)

[1.1 Развитие и особенности рынка ВДО 8](#_Toc134490418)

[1.1.1 Исторический обзор рынка ВДО 8](#_Toc134490419)

[1.1.2 Обзор современного рынка ВДО 9](#_Toc134490420)

[1.1.3 Развитие рынка ВДО в России 12](#_Toc134490421)

[1.2 Обзор деятельности кредитных рейтинговых агентств 16](#_Toc134490422)

[1.2.1 Кредитные рейтинговые агентства и их роль в развитии рынка облигаций 16](#_Toc134490423)

[1.2.2 Кредитные рейтинговые агентства в России 21](#_Toc134490424)

[1.3 Методы моделирования кредитного рейтинга 24](#_Toc134490425)

[1.3.1 «Классические» подходы к моделированию кредитного рейтинга эмитентов облигаций 24](#_Toc134490426)

[1.3.2 Применение методов машинного обучения для моделирования кредитных рейтингов эмитентов облигаций 25](#_Toc134490427)

[1.3.3 Подходы оценки качества предсказательных моделей 28](#_Toc134490428)

[1.3.4 Особенности моделирования кредитного рейтинга российских эмитентов ВДО 29](#_Toc134490429)

[1.4 Постановка гипотез 31](#_Toc134490430)

[Глава II. Построение моделей и интерпретация результатов 33](#_Toc134490431)

[2.1 Сбор данных 33](#_Toc134490432)

[2.3 Используемые модели предсказания кредитных рейтингов и методы оценки их качества 37](#_Toc134490433)

[2.5 Дескриптивный и корреляционный анализ 42](#_Toc134490434)

[2.7 Результаты моделирования кредитных рейтингов 49](#_Toc134490435)

[2.7.1 Подбор гиперпараметров 49](#_Toc134490436)

[2.7.2 Оценка качества предсказательных моделей 51](#_Toc134490437)

[2.8 Выводы 57](#_Toc134490438)

[Список литературы 61](#_Toc134490439)

[Приложение 67](#_Toc134490440)

# Введение

Предпосылки проведения исследования.Текущий этап развития российского финансового рынка характеризуется выходом на него ряда малых игроков: частных инвесторов, а также сравнительно небольших и молодых компаний. Малые и средние предприятия начинают выпускать облигации с высокой доходностью, которые привлекают частных инвесторов, стремящихся эффективнее распорядиться собственными сбережениями. Для оценки рисков эмитентов частные инвесторы в основном используют кредитные рейтинги. Следовательно, кредитные рейтинги оказывают решающее значение при покупке тех или иных облигаций. Однако ввиду того, что эти тенденции на российском финансовом рынке стали проявляться совсем недавно, научная литература, посвященная кредитным рейтингам российских эмитентов высокодоходных облигаций, сильно ограничена.

*Актуальность.* Рынок высокодоходных облигаций на российском финансовом рынке сформировался совсем недавно, а именно в 2018 году, хотя на рынках США и Европы рынок ВДО сформировался ещё во второй половине 20 века. Конечно, на российском рынке были эмитенты, которые и до 2018 года размещали свои облигации под сравнительно большие проценты, однако это явление не было массовым. Сейчас же рынок ВДО подает большие надежды. Компании решают выпустить высокодоходные облигации, потому что для них привлечение средств таким способом зачастую проще и дешевле, чем банковский займ.

Одной из наиболее значимых характеристик эмитента, указывающей на его финансовое состояние является кредитный рейтинг. Чем выше кредитный рейтинг эмитента, тем стабильнее его финансовое положение, а следовательно, тем ниже он сможет установить процентную ставку по своим облигациям. Частные инвесторы, которые являются основными покупателями высокодоходных облигаций, часто не обладают достаточными финансовыми компетенциями и специальным инструментарием, чтобы оценивать финансовое положение эмитентов, поэтому большинство из них руководствуется именно кредитными рейтингами при принятии решении о покупке тех или иных облигаций.

При этом получение кредитного рейтинга является дорогостоящей процедурой для эмитентов, что особенно значимо для тех, чье финансовое положение неустойчиво. Это не позволяет компаниям часто обновлять свои кредитные рейтинги, хотя при этом их финансовое состояние может резко меняться, что может привести к быстрому устареванию недавно полученных кредитных рейтингов. Помимо этого, в настоящее время получение кредитного рейтинга является обязательным условием для размещения облигаций на Московской бирже. Эти факты указывают на то, что на данный момент моделирование кредитных рейтингов российских эмитентов облигаций высокодоходных является актуальной задачей.

*Профессиональная значимость.* Результаты данного исследования могут быть полезны частным инвесторам для регулярной переоценки финансового состояния эмитентов, чьи облигации включены в их инвестиционные портфели, а также при приобретении облигаций в случаях, когда кредитный рейтинг был получен компанией относительно давно. Помимо инвесторов выводами данного исследования могут пользоваться и эмитенты облигаций, для формирования корректных ожиданий о том, какой кредитный рейтинг может быть присвоен им в ближайшем будущем.

*Гипотезы.* В рамках данного исследования рассматриваются следующие гипотезы:

* Модели, основанные на логистической регрессии, обладают адекватной предсказательной способностью.
* Применение методов машинного обучения позволяет значительно повысить качество прогнозирования кредитного рейтинга эмитентов облигаций по сравнению с логистической регрессией.
* Выявление нелинейных связей позволяет повысить предсказательную способность модели.
* Включение нефинансовых и макроэкономических показателей в модель повышает её предсказательную способность.

*Цели.* Данное исследование преследует цель построения и сопоставления различных моделей, предсказывающих кредитный рейтинг российских эмитентов высокодоходных облигаций на основе информации, находящейся в открытом доступе.

*Задачи.* Для достижения поставленной цели необходимо решить следующие задачи:

1. Дать определение термину «высокодоходные облигации», проанализировать особенности рынка ВДО и роль кредитных рейтинговых агентств в его развитии
2. Проанализировать научную и профессиональную литературу об основных способах и подходах к моделированию кредитных рейтингов эмитентов облигаций
3. Выделить особенности моделирования кредитного рейтинга эмитентов низкого кредитного качества по сравнению с эмитентами высокого кредитного качества
4. Изучить наиболее релевантные финансовые, нефинансовые и макроэкономические показатели, используемые для оценки кредитного рейтинга эмитентов высокодоходных облигаций
5. Собрать и обработать данные, необходимые для проведения исследования
6. Построить модели предсказания кредитного рейтинга, оценить их качество и выбрать среди них наиболее релевантную модель
7. Интерпретировать результаты и сделать выводы относительно поставленных гипотез

*Объект исследования.* Объектом данного исследования являются российские эмитенты высокодоходных облигаций, чьи облигации размешены на Московской бирже.

*Предмет исследования.* Кредитный рейтинг российских эмитентов высокодоходных облигаций, полученный с помощью статистических моделей, основанных на анализе финансовых и нефинансовых показателей эмитента, а также макроэкономических факторов.

*Информационная база исследования.* В рамках данного исследования использовались следующие базы данных: Cbonds (информация о кредитных рейтингах и финансовая отчетность эмитентов), СПАРК (нефинансовые показатели эмитентов), Росстат (макроэкономические показатели). Данные актуальны по состоянию на 28.02.2023.

*Степень проработанности проблемы.* В настоящий момент времени существует достаточно большой объем научной литературы, посвященной моделированию кредитных рейтингов эмитентов облигаций. Однако практически все научные работы по данной проблематике посвящены рынку облигаций инвестиционного кредитного качества. Те исследователи, которые включали в свои исследования высокодоходные облигации, как правило, не рассматривали их отдельно от облигаций инвестиционного качества. Однако некоторые исследования указывали на то, что моделирование высокодоходных облигаций может иметь ряд особенностей. Также подавляющее большинство исследований, посвященных моделированию кредитных рейтингов эмитентов облигаций проведены на примерах развитых рынков облигаций США и стран Европейского Союза, только в последнее время начали появляться аналогичные исследования для финансовых рынков стран Азии: Китая, Японии, Южной Кореи. Таким образом, российский рынок высокодоходных облигаций является малоизученным. При этом его изучение является важной исследовательской задачей, потому что российский рынок может иметь особенности по сравнению с финансовыми рынками других стран. А сам рынок высокодоходных облигаций в свою очередь становится все более значимым для российской экономики.

# Глава I. Теоретическая база исследования и обзор рынка

## 1.1 Развитие и особенности рынка ВДО

### 1.1.1 Исторический обзор рынка ВДО

В 70-х годах в США начался быстрый рост рынка высокодоходных облигаций, которые также нередко называют «мусорными облигациями». Высокодоходные облигации – это облигации, по которым выплачиваются более высокие процентные ставки, так как их эмитенты обладают более низким кредитным качеством, чем по облигациям финансово-устойчивых эмитентов (Investopedia, 2022). По началу это были облигации, так называемых «падших ангелов», компаний, чье финансовое положение резко ухудшилось, и они были вынуждены выпускать облигации с более высокой процентной ставкой, чем компании, находящиеся в лучшем финансовом положении. Позже на рынок высокодоходных облигаций начали выходить и молодые компании, подававшие большие надежды, но для которых получение банковского займа было слишком дорогой формой привлечения денежных средств, их начали называть «восходящими звездами». Высокая доходность таких облигаций начала привлекать инвесторов, которые стали активно их скупать, постепенно забывая о том, что высокая доходность этих облигаций была сопряжена с высокими рисками дефолта эмитентов. Следует отметить, что в тот период времени был достаточно низкий уровень дефолтов среди эмитентов ВДО, обусловленный благоприятной экономической ситуацией в США в 70-х годах двадцатого века.

Эйфория, связанная с бумом на рынке высокодоходных облигаций, закончилась в 1980-х годах, когда прошла волна дефолтов, вызванная рядом неблагоприятных макроэкономических факторов. Но уже в 1990-х годах рынок высокодоходных облигаций вновь начал расти, хотя уже более скромными темпами (Altman & Simon, 1994). К тому моменту инвесторы уже лучше осознавали риски высокодоходных облигаций.

### 1.1.2 Обзор современного рынка ВДО

В XXI веке рынок высокодоходных облигаций в значительной мере пострадал от мирового финансового кризиса 2008-2009 года. Это объясняется тем, что в условиях финансовой нестабильности, инвесторы прежде всего ожидают дефолтов от наименее финансово устойчивых компаний, к которым относятся эмитенты высокодоходных облигаций. Поэтому инвесторы старались выйти из рискованных активов в пользу наиболее безрисковых: государственных облигаций развитых стран (США, Великобритания и так далее) и облигаций инвестиционного класса. Что и произошло с рынком высокодоходных облигаций во время мирового финансового кризиса 2008-2009 года.

Однако после кризиса 2008-2009 рынок высокодоходных облигаций начал резко расти несмотря на то, что волна дефолтов эмитентов ВДО была достаточно большой. Это связано с тем, что после кризиса 2008-2009 года началась политика снижения процентных ставок по всему миру, особенно в развитых странах. Это привело к тому, что процентные ставки по государственным облигациям развитых стран начали сильно снижаться, в свою очередь процентные ставки по корпоративным облигациям инвестиционного класса также начали снижаться вслед за ними. Это привело к тому, что высокодоходные облигации вновь стали интересны инвесторам, так как они обладали процентными ставками значительно большими по сравнению с остальными облигациями.

При этом небольшим компаниям после кризиса 2008-2009 стало тяжелее получать банковские займы ввиду значительно ужесточившегося банковского регулирования, что очень сильно влияло на политику выдачи займов банков по отношению к компаниям с неустойчивым финансовым положением. Это привело к тому, что большое число таких компаний начали выпускать собственные облигации. Таким образом, этот факт в совокупности с растущим интересом со стороны инвесторов к высокодоходным облигациям обусловило начало «золотой эры» рынка высокодоходных облигаций в 2010 году, которая продлилась до 2015 года (Rosario, 2015). Стремительное развитие рынка высокодоходных облигаций привело к тому, что все большее исследователей начали посвящать свои работы анализу этого рынка и инвестиционной привлекательности такого финансового инструмента.

Так, исследователи выяснили, что динамика рынка высокодоходных облигаций очень сильно отличается от динамики рынка облигаций эмитентов высокого кредитного качества. Как было описано ранее на примере финансового кризиса 2008-2009, рынок высокодоходных облигаций в условиях кризиса начинает очень сильно падать, а рынок низкодоходных облигаций наоборот укрепляется. Если в условиях кризиса динамика рынка высокодоходных облигаций является предсказуемой, то динамика этого рынка в периоды экономического спокойствия и роста является менее очевидной, так как высокодоходные облигации менее волатильны по сравнению с низкодоходными (Теплова & Родина, 2019). Однако такая динамика рынка высокодоходных облигаций также может быть объяснена: в процентную ставку высокодоходных облигаций в основном заложен специфический риск дефолта эмитента и поэтому изменение процентных ставок в экономике не так сильно повлияет цену высокодоходных облигаций, в то время как облигации инвестиционного класса очень чувствительны к изменениям процентных ставок, так как сама по себе процентная ставка низкодоходных облигаций близка к ставкам по безрисковым активам, потому что риск дефолта эмитента находится на очень низком уровне. То есть можно выделить следующее: рынок высокодоходных облигаций в основном подвержен специфическому риску его эмитентов, а низкодоходных облигаций – процентному риску.

Следующее важное наблюдений исследователей рынка высокодоходных облигаций связано с их реальным уровнем риска. Так, Vazza и др., (2019) установили, что уровень дефолтов по облигациям с кредитным рейтингом BB (согласно рейтинговой шкале S&P) во времена кризисов составляет только 4,5%, что является допустимым для включения таких облигаций даже в самые консервативные инвестиционные портфели. При этом доходность таких облигаций значительно выше, чем по облигациям ближайшего кредитного качества, являющегося инвестиционным (BBB- согласно рейтинговой шкале S&P), хотя уровень дефолтов по таким облигациям не различается так значительно (Карминский, 2015). Однако стоит отметить, что уровень дефолтов по облигациям с кредитным рейтингом на уровне B (согласно рейтинговой шкале S&P) и ниже начинает резко расти по мере снижения рейтинга, и в таких случаях анализ и прогнозирование доходности портфеля на основе исторических данных становится некорректным, так как возрастает вес текущего специфического риска дефолта эмитентов.

Помимо этого, важной особенностью рынка высокодоходных облигаций является тот факт, что эмитенты высокодоходных облигаций «резко» объявляют дефолты и соответственно их кредитный рейтинг может меняться очень быстро (Vazza и др., 2019). Эмитенты с наивысшим кредитным рейтингом в среднем объявляют дефолт через 18 лет, если они вообще объявит его. Эмитенты ВДО с кредитным рейтингом на уровне B в среднем объявят дефолт через 5 лет после размещения облигаций или вовсе не объявят его. Для эмитентов с кредитным рейтингом ниже ССС такой период сокращается до 2 лет. При этом разброс медианного периода между размещением облигаций и дефолтом эмитента со средним уровнем кредитного качества (на уровне BB- (согласно рейтинговой шкале Fitch)) и низким (на уровне CCC (согласно рейтинговой шкале Fitch)) составляет 7 лет и 3 месяца соответственно. Этот факт также указывает на то, что необходимо регулярно анализировать финансовое положение эмитентов высокодоходных облигаций, во избежание неожиданных дефолтов облигаций в инвестиционном портфеле.

Также важной особенностью рынка высокодоходных облигаций является его сравнительно низкая ликвидность относительно облигаций с высоким кредитным рейтингом (SEC, 2013). Так как высокодоходные облигации остаются нишевым финансовым инструментом, то данные облигации обладают достаточно низкой ликвидности, что в периоды финансовой неустойчивости может привести к ситуации, когда выйти из такого актива может быть достаточно сложно, и потребует от их владельца продать облигации с большим дисконтом.

Все эти особенности рынка высокодоходных облигаций сводятся к тому, что инвестору, который решает инвестировать в высокодоходные облигации необходимо проводить частую оценку финансового состояния эмитентов облигаций, потому что события на рынке высокодоходных облигаций развиваются стремительно, а в случае не своевременной реакции на негативные события есть риск не иметь возможности продать облигации, по которым резко выросла вероятность дефолтов.

### 1.1.3 Развитие рынка ВДО в России

Рынок высокодоходных облигаций в России сформировался сравнительно недавно, а именно в 2018 году. При этом он смог сформироваться только в так называемую «вторую волну ВДО». Первая волна рынка ВДО закончилась в 2008-2009, когда множество эмитентов высокодоходных облигаций было вынуждено объявить себя банкротами. Вторая волна рынка ВДО началась только в 2016 году. Однако и в периоде после 2009 года и до 2016 года были эмитенты, чьи облигации имели доходность значительно выше, чем по остальным облигациям. Это были компании, испытывавшие значительные финансовые трудности, что вынудило их выпускать облигации с высокой процентной ставкой. Известными примерами таких компаний являются: «Трансаэро», «ЮТэйр», «Мечел» и так далее. Некоторые «падшие ангелы» смогли урегулировать свое финансовое положение, как например, «Мечел». Однако были и компании, как «Трансаэро», которые не смогли это сделать и объявили себя банкротами.

Вторая волна ВДО началась в 2016 году с выпуском облигаций молодой компанией «Ломбард Мастер». После чего на российский рынок начали активно выходить молодые эмитенты облигаций, предлагавшие высокие процентные ставки по своим облигациям. В основном это были финансовые компании (МФО и лизинговые компании), однако ближе к 2018 году на рынок начали выходить и нефинансовые компании, что в последствие начало свидетельствовать о формировании рынка высокодоходных облигаций.

В России с 2019 года начали действовать меры государственной поддержки эмитентов высокодоходных облигаций, которые с одной стороны были нацелены на субсидирование ставки купона, а с другой стороны –поддержки эмитентов при размещении облигаций. Эти меры свидетельствует о том, что российские органы государственной власти начали осознавать важность рынка высокодоходных облигаций для экономики в целом.

На данный момент времени не существует точного определения высокодоходных облигаций в России. В развитых странах под высокодоходными облигациями часто понимают облигации, чьи эмитенты (гаранты) обладают кредитным рейтингом международного кредитного агентства на уровне BB+ (согласно рейтинговой шкале S&P) и ниже (NASDAQ, 2018). Однако российские эмитенты в крайне редких случаях получают кредитные рейтинги от международных агентств, что требует разработки критериев определения высокодоходных облигаций для российского финансового рынка. При этом «первопроходцы» рынка высокодоходных облигаций в России не имели кредитных рейтингов вовсе. Однако, начиная с октября 2020 года, все эмитенты, желающие выпустить разместить свои облигации на Московской бирже обязаны получить кредитный рейтинг (Московская Биржа, б.д.). При условии, что подавляющее большинство эмитентов высокодоходных облигаций выбирают именно биржевое размещение своих облигаций.

Инвестиционные компании и аналитические агентства в основном используют следующие критерии для определения высокодоходных облигаций, разработанные компанией Cbonds для внутренних расчетов (Cbonds, б.д.):

1. Ставка купонной доходности облигации на момент установления превышает ключевую ставку не менее чем на 5% годовых.
2. Объем эмиссии не превышает 2 млрд рублей.
3. Срок обращения составляет не менее 360 дней.

С точки зрения анализа рисков российский рынок высокодоходных облигаций имеет ряд важных особенностей, которые необходимо учитывать при его анализе (Теплова & Родина, 2019):

1. *Низкая ликвидность.* Данная черта присуща в разной степени любому рынку высокодоходных облигаций, однако для российского рынка данная черта является особенно значимой, так как данный сегмент финансового рынка развит достаточно слабо и поэтому достаточно малое число инвесторов приобретает высокодоходные облигации, что в условиях кризиса может привести к тому, что выйти из данного актива можно будет только со значительными потерями в капитале.
2. *Низкое качество аналитической информации.* Несмотря на то, что существуют хорошо зарекомендовавшие себя методологии оценки финансового состояния компаний, множество из них не применимо для российских эмитентов высокодоходных облигаций, так как они разработаны для работы с отчётностью по МСФО. Большая часть российских эмитентов ВДО выпускают исключительно отчётность по РСБУ, при этом нередко с рядом ошибок (Семерина & Байбеков, 2014). Это не позволяет точно оценивать финансовое состояние эмитентов высокодоходных облигаций.
3. *Отсутствие достаточной истории дефолтов*. Ввиду того, что. Рынок ВДО сформировался только к 2018 году, не было еще ни одной значимой серии дефолтов эмитентов облигаций, которое не позволяет судить достоверно ли оцениваются риски российских эмитентов высокодоходных облигаций в настоящее время. Использование данных за 2008-2009 годы не является корректным, потому что российский рынок ВДО претерпел серьезные изменения с тех пор.
4. *Государственная поддержка МСП.* На данном этапе развития рынка высокодоходных облигаций государство оказывает значительную поддержку малым и средним предприятиям, что позволяет им быть более устойчивыми в условиях кризиса.
5. *Необходимость регулярной оценки финансового состояния эмитента.* На российском рынке высокодоходных облигаций большое число компаний объявляет дефолт уже при первых купонных выплатах, а также в периоды, крупных выплат. Данный факт указывает на то, что инвестору стоит оценивать финансовое состояние эмитентами на регулярной основе, особенно перед значимыми событиями.

В основном высокодоходные облигации ориентированы на частных инвесторов, так как большинство институциональных инвесторов, в особенности негосударственные пенсионные фонды, не могут покупать высокодоходные облигации ввиду высоких рисков эмитентов. Этот факт необходимо учитывать при анализе рынка высокодоходных облигаций в России, потому что частные инвесторы в основном не обладают достаточными финансовыми знаниями и компетенциями для всестороннего анализа финансового положения эмитентов, что приводит к тому, что кредитные рейтинги являются одним из главных ориентиров о покупке тех или иных облигаций инвестором (Карминский, 2015).

Причины стремительного развития рынка высокодоходных облигаций в России и в развитых странах схожи (Теплова & Родина, 2019):

1. Процентные ставки по высокодоходным облигациям значительно выше, чем по облигациям инвестиционного качества.
2. Для малых и средних компании выпуск высокодоходных облигаций является более дешевым способом финансирования собственной деятельности.
3. Привлечение средств путем выпуска высокодоходных облигаций позволяет диверсифицировать источники финансирования и сделать соответствующую политику эмитента более гибкой.

Таким образом, можно сделать выводы о том, что высокодоходные облигации являются перспективным финансовым инструментом, однако на российском рынке они появились совсем недавно, из-за чего данная тема еще не изучена в полной мере. Оценка рисков высокодоходных облигаций является ключевой задачей при приобретении тех или иных ВДО. А ввиду быстрой изменчивости рынка высокодоходных облигаций такие оценки необходимо проводить часто. Так как главными покупателями высокодоходных облигаций являются частные инвесторы, то анализ высокодоходных облигаций тесно связан с кредитными рейтингами, на которые ориентируется большая часть инвесторов.

## 1.2 Обзор деятельности кредитных рейтинговых агентств

### 1.2.1 Кредитные рейтинговые агентства и их роль в развитии рынка облигаций

После того, как инвесторы столкнулись с сериями дефолтов эмитентов облигаций, инвесторы осознали необходимость оценки финансового состояния эмитента при принятии решения о покупке выпущенных им облигаций. Однако большинство частных инвесторов не обладало достаточными знаниями и компетенциями для проведения качественного анализа финансового состояния эмитентов облигаций, что порождало спрос на то, чтобы сторонняя организация независимо оценивала финансовое положение эмитентов облигаций.

Так, ближе к концу 90-х годов независимые кредитные рейтинговые агентств, появившиеся еще в начале 20 века, начинали играть все более важную роль на рынке облигаций в США, а вскоре начали играть значительную роль на финансовых рынках других стран. Кредитное рейтинговое агентство – это организация, которая оценивает кредитоспособность и финансовую устойчивость компаний, правительств и других организаций, выпускающих облигации и другие ценные бумаги (Cbonds, б.д.). Кредитные рейтинговые агентства присваивают организациям кредитные рейтинги, под которыми следует понимать «мнения кредитных рейтинговых агентств о способности организации исполнять принятые на себя финансовые обязательства (кредитоспособность, финансовая надежность, финансовая устойчивость) и (или) о кредитном риске ее отдельных финансовых обязательств или финансовых инструментов» (Банк России, 2023).

Рейтинговые агентства присваивают кредитные рейтинги эмитентам облигаций на основе всестороннего анализа их финансового состояния. Для этого они используют различные подходы, методы, которые можно объединить в следующие группы (НКР, 2023; Эксперт Ра, 2023; НРА, 2023; АКРА, 2023):

1. *Финансовые показатели компании.* Рейтинговые агентства рассчитывают большое число финансовых показателей на основе бухгалтерской отчётности компаний, которые характеризуют ликвидность активов, долговую нагрузку, способность компании обслуживать текущие обязательства, рентабельность компании и так далее.
2. *Положение компании на рынке и состояние отрасли в целом.* Важным фактором при оценке кредитного рейтинга компании является её способность конкурировать с другими компаниями в отрасли и получать прибыль в своей отрасли. Так, рейтинговые агентства тщательно анализируют долю компании на рынке, ее сравнительные преимущества над конкурентами, а также ее потенциал к росту и развитию.
3. *Риск-менеджмент и качество корпоративного управления.* Также рейтинговые компании оценивают на сколько профессионально и прозрачно управляется компания. Ввиду того, что рейтинговое агентство оценивает перспективы компании в ближайшем будущем, они учитывают эффективность менеджмента и его способность принимать эффективные решения даже в неблагоприятной экономической ситуации.
4. *Макроэкономические факторы.* Рейтинговые агентства учитывают различные макроэкономические показатели (инфляцию, процентные ставки, рост ВВП и так далее), которые могут оказывать сильное влияние на финансовую устойчивость компании и ее потенциал к росту.

Исходя из проведенного анализа рейтинговое агентство присваивают компаниям кредитный рейтинг. Этот кредитный рейтинг не является бессрочным, и компаниям необходимо обновлять его регулярно. Так, например, Московская биржа требует от эмитентов облигаций обновлять кредитный рейтинг не реже, чем раз в год (Московская Биржа, б.д.).

Стоит отметить, что кредитный рейтинг — это прежде всего субъективное мнение агентства. Это приводит к тому, что для рейтингового агентства важен уровень доверия к нему, которое формируется в течение долгого периода времени. Именно поэтому рейтинговые агентства обязаны демонстрировать свою независимость при присвоении кредитных рейтингов (Maher & Sen, 1997). В данном случае под независимостью стоит понимать то, что никакая из сторон: сами эмитенты, регуляторы и иные третьи лица не имеют возможности оказывать влияние на принятие решений кредитным рейтинговым агентствам. Несмотря на субъективность кредитных рейтингов, рейтинговые агентства должны использовать объективные методы и критерии для оценки финансовой устойчивости компании, а также сам процесс должен быть максимально прозрачным (Maher & Sen, 1997). Именно поэтому все крупные кредитные рейтинговые агентства публикуют подробные методологии того, как они оценивают компании. Несмотря на то, что каждое рейтинговое агентства использует достаточно похожие критерии оценки, сам набор оцениваемых показателей и их вес при принятии окончательного решения может очень сильно отличаться у различных рейтинговых агентств, что приводит к тому, что кредитный рейтинг компании может зависеть от того, какое рейтинговое агентства оценивало компанию, даже если речь идет о крупнейших и наиболее авторитетных рейтинговых агентствах (Daniel и др., 2019). Именно поэтому существуют шкалы, которые демонстрируют, как соотносятся кредитные рейтинги разных рейтинговых агентств. Строятся эти шкалы на основе анализа истории дефолтов компаний и кредитного рейтинга, который они имели.

Достаточно сложно переоценить роль кредитных рейтинговых агентств в развитии рынка облигаций в особенности ВДО. Благодаря тому, что компании получают кредитные рейтинги ряд частных инвесторов, смогли выйти на финансовый рынок, так как оценка финансового состояния компании является достаточно сложным процессом, зачастую требующий наличия ряда профессиональных компетенций. На основе кредитных рейтингов инвесторы могут взвешивать доходность облигации и риски, которая они несут, а также сопоставлять облигации разных эмитентов между собой. Данный факт наиболее актуален для рынка высокодоходных облигаций, потому что основными покупателями таких облигаций являются частные инвесторы, а не институциональные инвесторы, которые зачастую не имеют юридической возможности покупать ценные бумаги ненадёжных эмитентов (Теплова & Родина, 2019). Однако будет ошибочным полагать, что только частные инвесторы пользуются кредитными рейтингами. Регуляторы разных стран и различные фонды также ориентируются на кредитный рейтинг. Так, например, различные фонды декларируют свою стратегию и уровень риска, которому подвержены, и поэтому они могут прописывать требования по кредитному рейтингу эмитентов облигаций, которые приобретают. Регуляторы тоже могут прописывать требования по кредитному рейтингу эмитентов ценных бумаг, которые покупают фонды, особенно это актуально для частных пенсионных фондов, уровень риска, которых должен быть минимальным.

Для эмитентов облигаций очень важно, какой кредитный рейтинг они получат, так как он будет оказывать прямое влияния на то, какую процентную ставку они будут вынуждены предложить своим инвесторам. При этом даже небольшое различие в кредитном рейтинге может значительно повлиять на то, какую ставку эмитент будет вынужден предложить своим инвесторам, что может повлиять на его финансовую устойчивость в будущем. Так, спред доходности для облигаций уровня Baa3 и Ba1 (согласно шкале рейтингового агентства Moody’s) может отличаться на 100 базисных пунктов (Карминский, 2015) (Maher & Sen, 1997), хотя это соседствующие уровни рейтинга. Но при этом облигации эмитентов с кредитным рейтингом Baa3 являются облигациями инвестиционного уровня, когда облигации уровня Ba1 уже считаются облигациями спекулятивного кредитного качества, из-за чего их не могут приобретать некоторые институциональные инвесторы, в том числе негосударственный пенсионные фонды. При этом существует исследования, которые подтверждают тот факт, что присвоение кредитных рейтингов может нарушать эффективное образование доходности облигаций (Kräussl, 2003). Так в случаях, когда у компании кредитного рейтинга нет, процентная ставка формируется на справедливом уровне, а получение кредитного рейтинга может нарушать это состояние, то есть имеет место непосредственное влияние кредитного рейтинга на спред доходности облигации при прочих равных. Но при этом само по себе получение кредитного рейтинга повышает уровень доверия к компании, что в свою очередь понижает требуемую процентную ставку, а также при получении кредитного рейтинга растёт число упоминаний компании в средствах массовой информации, что положительно влияет на успешность размещения облигаций (Карминский, 2015).

Несмотря на то, что кредитные рейтинги в значительной мере оказывают благоприятное влияние на развитие рынка облигаций в целом, для малых и средних предприятий получение кредитного рейтинга является очень дорогой процедурой. Высокая цена услуг кредитных рейтинговых агентства обусловлена тем, что для проведения всестороннего анализа финансовой устойчивости компаний, рейтинговое агентство должно обработать большой объем информации, как о самой компании, так и об отрасли в целом. Ввиду высокой значимости деятельности кредитных рейтинговых агентств их ошибка может нести очень большие репутационные издержки, что требует от экспертов агентства высокого уровня профессионализма.

Однако кредитные рейтинговые агентства нередко подвергаются критике за то, что их кредитные рейтинги не всегда отражают реальную финансовую устойчивость компаний. Так, например, в 2008-2009 году значительная часть обанкротившихся компаний обладали наивысшим кредитным рейтингом (Теплова & Родина, 2019). Помимо этого, в последнее время наблюдается «инфляция» кредитных рейтингов, при которой рейтинговые агентства с каждым годом всё чаще присваивают высокие кредитные рейтинги эмитентам облигаций, хотя явных оснований к этому может и не быть (Goldstein & Chong, 2020). Эти факты указывают на то, что к кредитным рейтингам необходимо относиться достаточно критично, но несмотря на это кредитные рейтинги продолжают оставаться значимыми для финансового рынка любой страны.

### 1.2.2 Кредитные рейтинговые агентства в России

Одним из первых российским кредитным рейтинговым агентством (КРА) является «Эксперт РА», который был основан в 1997 году. Со временем начали появляться и другие значимые рейтинговые агентства: «АКРА», «Национальное Рейтинговое Агентство» (НРА) и «Национальные Кредитные Рейтинги» (НКР)», «РусРейтинг» и «Рейтинговое агентство АК&M». В России, как и в ряде других стран рейтинговые агентства начали играть важную роль, что потребовало со стороны государственных регуляторов проводить аккредитацию кредитных рейтинговых агентств. Так, на данный момент в реестр кредитных рейтинговых агентств входят 4 компании: «АКРА», «Эксперт РА», «Национальное Рейтинговое Агентство» (НРА) и «Национальные Кредитные Рейтинги» (НКР)» (Банк России, 2023). Рейтинги этих 4 компаний являются наиболее авторитетными. Кредитные рейтинги используются не только инвесторами и эмитентами, но и Московской Биржей при включении облигаций в котировальные списки I и II уровней, страховыми организациями, банками при оценке рисков своих контрагентов и заемщиков, а также негосударственными пенсионными фондами при покупке ценных бумаг различных компаний, а также некоторыми юридическими лицами при оценки рисков тех или ценных бумаг для формирования соответствующих резервов (Банк России, 2023).

Эмитенты обращаются к услугам российских рейтинговых агентств, а не международных, потому что цены на услуги отечественных КРА значительно ниже, а также российские кредитные рейтинговые агентства лучше осведомлены об экономической ситуации в стране и отрали, что позволяет давать компаниям более точную оценку их финансового состояния. Поэтому если компания не имеет намерения выходить на международный финансовый рынок, то ей будет рациональнее получить кредитный рейтинг у российского кредитного рейтингового агентства.

Однако на данный момент времени сопоставление кредитных рейтингов различных рейтинговых агентств является достаточно неточным, при этом, какое кредитное рейтинговое агентство оценивало эмитента, может влиять на то, какой кредитный рейтинг будет в итоге присвоен (Daniel и др., 2019). Российские эмитенты начали массово получать кредитные рейтинги не так давно, и при этом их получали в основном крупные компании, которые и так обладали высоким уровнем финансовой устойчивости. На данный момент времени не существует достаточно продолжительной истории дефолтов компаний с кредитным рейтингом, что не позволяет точно сопоставить кредитные рейтинги различных рейтинговых агентств. Поэтому Центральный Банк Российской Федерации (Банк России, б.д.) предоставляет свою методологию сопоставления кредитных рейтингов различных кредитных рейтинговых агентств, в которых уровни кредитного рейтинга объединены в группы по 1-2 уровня (Банк России, 2021), она представлена в приложении 1.

Таким образом, кредитные рейтинги играют важную роль в функционировании и развитии рынка облигаций. Благодаря им все больше частных инвесторов выходит на финансовый рынок, что является необходимым условием для развития рынка высокодоходных облигаций. Несмотря на субъективность кредитных рейтингов, участники фондовых рынков по всему миру продолжают доверять рейтинговым агентствам, принимая решения о покупке тех или иных ценных бумаг на основе их мнения. Однако остается нерешенной проблема высокой стоимости услуг получения кредитного рейтинга, что является особо острой проблемой для малых и средних компаний, которые являются основными эмитентами высокодоходных облигаций. Отказ от получения кредитного рейтинга может негативно сказаться на успешности размещения облигаций, так как инвесторы могут с недоверием относиться к таким эмитентам, а также ряд фондовых бирж, в том числе Московская биржа, требует от эмитентов наличие кредитного рейтинга для размещения их облигаций на бирже. Нерешенность этих проблем порождает интерес среди исследователей к моделированию кредитных рейтингов, которое позволило бы инвесторам и эмитентам оценивать кредитный рейтинг на регулярной основе и с минимальными затратами.

## 1.3 Методы моделирования кредитного рейтинга

### 1.3.1 «Классические» подходы к моделированию кредитного рейтинга эмитентов облигаций

Одними из первых работ в области моделирования кредитного рейтинга эмитентов облигаций начали появляться во второй половине 20 века и были посвящены американскому рынку облигаций нефинансовых компаний.

Наиболее значимыми являются работы Fisher’a (1959), Horrigan’a (1996) и Pogue’a & Soldofsky’a (1996). В них исследователи используют строят модели исключительно на основе финансовых показателей, которые можно объединить в следующие группы: ликвидность и оборачиваемость активов, долговая нагрузка, рентабельность компании. Также важной чертой ранних исследований является то, что первые модели, предсказывающие кредитный рейтинг эмитентов облигаций, строились на основе логистической регрессии. Однако достаточно скоро исследователи начали применять и другие статистические методы. Так, например, (Pinches & Mingo, 1975) построили аналогичные модели на основе мультиклассового дискриминантного анализа, что в итоге позволило немного повысить качество предсказания, однако в последующих исследованиях данный метод использовался достаточно редко ввиду малой разницы в качестве предсказания в сравнении с логит-моделью.

Несмотря на то, что модели исследователей того времени были достаточно простыми: включали в себя исключительно финансовые показатели в качестве независимых переменных, а основными методами являлась логит-модель и дискриминантный анализ, эти модели обладали неплохим качеством предсказания, точность предсказания составляла около 55-65% (Maher & Sen, 1997).

Несмотря на невысокую предсказательную способность «классических подходов» исследователи и в настоящее время продолжают использовать их. Это объясняется тем, что логит-модель в отличие от более современных методов предсказания рейтингов является хорошо интерпретируемой. Это позволяет исследователям не только строить модели, способные предсказывать кредитные рейтинги, но и иметь представления, какие факторы являются наиболее значимыми, что позволяет им выделять новые и неочевидные связи, необходимые для точной оценки кредитных рейтингов компаний. На основе логит-моделей исследователи доказали, что кредитный рейтинг эмитентов облигаций может быть достаточно точно предсказан исключительно на финансовых показателях (Maher & Sen, 1997). Однако последние исследования указывают на то, что включения макроэкономических показатели оказывает значимое влияние на кредитный рейтинг, особенно для развивающихся стран (Karminsky и др. 2019). Помимо этого, исследователи обнаружили, что для предсказания кредитных рейтингов эмитентов облигаций США размеры компании играют большую роль, чем показатели отражающую долговую нагрузку компаний, а так как большинство первых исследований были посвящены именно американскому рынку, то методологии моделирования кредитных рейтингов эмитентов облигаций в большей степени опирались на показатели размеров компании и ее активов, при этом для эмитентов облигаций стран Евросоюза такой подход является неправильным и смещение фокуса в сторону характеристик долга и ликвидности активов способно повысить качество предсказания (Zhong и др., 2014).

Однако кардинально повысить качество предсказания «классических» моделей у исследователей не получалось путем включения новых переменных. Это побудило их искать другие методы для моделирования кредитного рейтинга, которые смогли бы повысить качество предсказания.

### 1.3.2 Применение методов машинного обучения для моделирования кредитных рейтингов эмитентов облигаций

Методы машинного обучения начали активно развиваться во второй половине 20 века. Исследователи проявляют большой интерес к таким методам, потому что они зачастую имеют большую предсказательную способность (Gu и др., 2020). Под методами машинного обучения стоит понимать класс методов искусственного интеллекта, для решения различных задач, на основе уже решённых аналогичных задач (Rudin, 2019).

После того, как большое число исследователей столкнулись с тем, что «классические методы» не позволяли получить очень высокое качество предсказания, они начали применять различные методы машинного обучения для предсказания кредитных рейтингов. Первые такие работы начали появляться в конце 90-х годов. Тогда исследователи начали применять нейронные сети (ANN). Модели в более ранних исследователи основывались исключительно на финансовых показателях, в основном использовались те же переменные, что и в более ранних исследованиях с моделями, основанными на логистической регрессии. Однако точность предсказания моделей, основанных на нейронных сетях, выросла до 86% (Maher & Sen, 1997). Такой высокое качество предсказания исследователи смогли повторить и на выборках других стран, в основном США, стран Европейского Союза и Тайваня. Так, качество моделей, основанных на нейронных сетях, составляло от 63% до 86%, при этом высокое качество предсказания демонстрировали компании, как с большим числом переменных (около 100), так и с значительно меньшим (от 5 до 10) (Parisa, 2020). Несмотря на то, что исследователи смогли добиться высокой точности предсказания кредитного рейтинга, они столкнулись с тем, что эти модели являлись совершенно не интерпретируемыми, что не позволяло оценить какие факторы оказывают наибольшее влияние на кредитный рейтинг эмитентов облигаций. Это связано с тем, что методы машинного обучения по своей сути являются «black box» (Rudin, 2019), то есть зачастую нельзя понять, каким образом модель предсказала то или иное значение, особенно данное суждение справедливо для нейронных сетей, которые способны комбинировать различные переменные с целью создания новых признаки, что полностью исключает интерпретацию признаков. Поэтому исследователи начали использовать другие методы машинного обучения, чьей алгоритм решения более интуитивно понятен.

Так, наиболее часто используемыми методами оказались: метод ближайших соседей (k-nearest neighbors), метод опорных векторов (support vector machine), решающее дерево (decision tree), случайный лес (random forest). В основе всех этих методах лежат совершенно разные подходы к определению кредитного рейтинга. Это может приводить к тому, что качество предсказания кредитного рейтинга может отличаться между различными методами.

Наибольшей предсказательной способностью обладают следующие методы: SVM (метод опорных векторов) и DT (решающее дерево) (Parisa, 2020). Точность данных методов составляет до 89% и 86% соответственно. При этом применение метода опорных векторов, как правило, позволяет получить более стабильное качество предсказания (от 60 до 89%), в то время как точность предсказания моделей, основанных на решающем дереве, составляет от 47 до 86%. Остальные методы, как правило, показывают меньшую точность предсказания, однако в некоторых случаях она может быть тоже высокой, что говорит о необходимости использования всех основных методов для прогнозирования кредитного рейтинга.

Исследователи особое внимание уделяют методу опорных векторов. Это связано с тем, что данный метод, как правило, имеет наибольшее качество предсказания и он может искать как линейные, так и нелинейные связи, в зависимости от выбранного ядра. При этом есть свидетельства того, что применение нелинейных ядер потенциально может повысить предсказательную способность модели (Ha ́jek & Olej, 2011).

Сопоставимым с методом опорных векторов по частоте использования является решающее дерево. В некоторых исследованиях утверждается, что данный метод является наиболее универсальным (Jiexian и др., 2021). Это объясняется тем, что индукция правил, которая лежит в основе данного метода, позволяет эффективно выделять, как линейные, так и нелинейные связи показателей эмитента с его кредитным рейтингом. При этом метод решающего дерева лежит в основе другого эффективного метода: случайного леса. По мнению некоторых исследователей (Wei, 2022), данный метод устойчив к излишней подстройке модели к обучающим данным, и поэтому сама модель будет обладать большей предсказательной силой на различных выборках.

Помимо этого, стоит отметить, что методы машинного обучения могут требовать особой предобработки данных (Hsu-Che и др., 2014). Так, помимо удаления выбросов, данные необходимо масштабировать, чтобы сам масштаб каждого признака не влиял итоговую модель (некоторые методы неустойчивы к этому, особенно KNN).

Таким образом, методы машинного обучения позволили получать модели, обладающие более высокой предсказательной способностью. Благодаря высокой точности предсказания такие модели могут эффективно применяться на практике, в отличие от моделей, построенных на основе «классических методов». Однако ввиду сложности интерпретации методов машинного обучения при проведении исследований, посвященных моделированию кредитного рейтинга, стоит совмещать, как «традиционные», так и «продвинутые» методы для более всестороннего анализа, полученных моделей.

### 1.3.3 Подходы к оценке качества предсказаний моделей

При сопоставлении различных методов для предсказания кредитных рейтингов эмитентов облигации исследователи прибегают к такой метрике, как точность предсказания. Однако такая метрика имеет некоторые недостатки (Kürüm и др., 2011):

1. Чувствительна к равномерности классов
2. Не учитывает неравнозначность ошибок первого и второго рода

Как правило, распределение кредитных рейтингов между компаниями является неравномерной величиной (Parisa, 2020). Это может приводить к тому, что предсказание наиболее часто встречающихся классов может приводить к завышению точности предсказания. Помимо этого, завышение и занижение кредитного рейтинга компании не является соразмерной ошибкой. Завышение кредитного рейтинга может привести к излишнему оптимизму инвесторов и эмитентов, что в случае наступления неблагоприятных экономических условий может привести к неожиданно высокому риску дефолта эмитента. Занижение же рейтинга скорее приведет к небольшим финансовым потерям либо исключительно утрате части прибыли. Поэтому для предсказания кредитного рейтинга следует использовать также следующие метрики качества предсказаний моделей: ROC-AUC и F1-score (Wang и др., 2020). Они позволяют учитывать несоразмерность ошибок первого и второго рода (учитывают баланс между «чувствительностью» и «специфичностью» предсказания). Помимо этого, к данным метрикам применяют методы усреднения: макроусреднение и микроусреднение, благодаря которым можно оценивать качество моделей на несбалансированных выборках. Включение этих метрик позволяет преодолеть вышеуказанные недостатки метрики «точность». Однако метрика «точность» является наиболее интуитивно понятной, поэтому стоит использовать все эти метрики в совокупности.

### 1.3.4 Особенности моделирования кредитного рейтинга российских эмитентов ВДО

Большая часть исследований о моделировании кредитных рейтингов эмитентов облигаций, посвящена рынкам облигаций инвестиционного качества развитых стран (Parisa, 2020). Данный факт может приводить к тому, что результаты данных исследования могут быть неприменимыми для рынков высокодоходных облигаций развивающихся стран ввиду их специфики. Так, исследования, посвященные финансовым рынкам конкретных стран, могут использовать значимые объясняющие переменные, применимые только для конкретной страны (Mirko и др., 2019).

Так, на примере рынков облигаций стран БРИКС (Karminsky и др., 2019) и Китая (Jing-Ping и др., 2020), было обнаружено, что макроэкономические факторы оказывают значимое влияние при предсказании кредитного рейтинга эмитентов. При этом макроэкономические факторы при моделирование кредитных рейтингов эмитентов из развитых стран используются редко (Cheng, 2020). Также есть свидетельства того, что различные факторы играют разную роль при предсказании различных уровней кредитного рейтинга. Так, при предсказании кредитного рейтинга эмитентов облигаций в Китае со спекулятивным кредитным рейтингом макроэкономические факторы играли большую роль по сравнению с эмитентами облигациями инвестиционного уровня. При этом точность предсказания кредитного рейтинга эмитентов спекулятивного кредитного качества было выше, чем эмитентов инвестиционного кредитного качества (Jing-Ping и др., 2020).

Помимо этого, основные подходы к моделированию кредитных рейтингов разрабатывались на основе международных стандартов финансовой отчетности, однако российские эмитенты высокодоходных облигаций в основном не предоставляют отёчность по МСФО, а только по РСБУ. При этом само качество отчетности российских эмитентов высокодоходных облигаций может содержать значительно число ошибок (Семерина & Байбеков, 2014), что может приводить к снижению предсказательной способности моделей.

При предсказании кредитных рейтингов эмитентов облигации исследователи часто используют показатели, связанные с ценой акций эмитентами (Теплова & Родина, 2019). Такие факторы являются значимыми и значительно повышают качество предсказания, однако в случае с российскими эмитентами высокодоходных облигаций их включение невозможно. Это связано с тем, что для российских эмитентов высокодоходных облигаций выпуск облигаций является первым шагом на пути к IPO, хотя в развитых странах компании обычно выходят на IPO, а потом выпускают облигации (Теплова & Родина, 2019).

Также при моделировании кредитного рейтинга эмитентов облигаций, как правило, качество менеджмента и внутреннего контроля не учитывается в моделях, хотя рейтинговые агентства оценивают их при присвоении кредитного рейтинга, однако в случае с высокодоходными облигациями данный фактор может иметь большое влияние на кредитный рейтинг (Семерина & Байбеков, 2014), так как качество управления в таких компаниях может значительно различаться. Помимо этого, математические модели не способны учитывать специфические черты каждой компании, которые могут оказывать влияние на кредитный рейтинг компании (Семерина & Байбеков, 2014).

## 1.4 Постановка гипотез

Учитывая описанное ранее, логистическая регрессия является «традиционным» хорошо зарекомендовавшим методом для предсказания кредитного рейтинга. Однако предсказательная способность модели нередко зависит от особенностей выборки (страны эмитентов, отрасли и так далее) на которой используется (Parisa, 2020). Это говорит о необходимости проверки предсказательной способности логистической регрессии на выборке российских эмитентов ВДО. Так, гипотеза 1 может быть сформулирована следующим образом: Модели, основанные на логистической регрессии анализе обладают адекватной предсказательной способностью.

Методы машинного обучения уже продемонстрировали высокую предсказательную способность в задачах моделирования кредитных рейтингов, однако качество таких моделей зачастую было нестабильным, что может говорить о их низкой применимости на некоторых выборках. Поэтому гипотеза 2 может быть сформулирована следующим образом: применение методов машинного обучения позволяет значительно повысить качество прогнозирования кредитного рейтинга эмитентов облигаций по сравнению с логистической регрессией.

Существуют свидетельства того, что влияние факторы определяющих кредитный рейтинг может быть нелинейным (Ha ́jek & Olej, 2011), что может говорить о необходимости использования нелинейных методов для моделирования кредитных рейтингов. Таким образом, гипотеза 3 может быть сформулирована следующим образом: выявление нелинейных связей позволяет повысить предсказательную способность модели.

Несмотря на то, что большинство исследований фокусируется исключительно на финансовых показателях (Cheng, 2020), макроэкономические показатели на некоторых выборках демонстрировали значимую связь (Jing-Ping и др., 2020), что потенциально может повысить качество предсказаний моделей. Включение нефинансовых и макроэкономических показателей в модель повышает её предсказательную способность.

# Глава II. Построение моделей и интерпретация результатов

## 2.1 Сбор данных

Выборка включает в себя 72 компании нефинансового сектора, обладающими кредитным рейтингом, чьи облигации включены в индекс Cbonds-CBI RU High Yield за период с 2018 по 2022 год (Cbonds, 2022). Размер выборки ограничен тем, что рынок высокодоходных облигаций является очень молодым, и поэтому рассмотреть большое количество компаний не представляется возможным в рамках данного исследования. Однако на примере ряда исследований было продемонстрировано, что малое число рассматриваемых компаний необязательно приводит к низкому качеству предсказания моделей. Так, Khemakhem & Boujelbene (2015) и Ye и др., (2008) смогли продемонстрировать, что выборки в менее 80 компаний все равно могут иметь высокую предсказательную способность. Стоит отметить, что среди эмитентов ВДО многие компании начали получать кредитный рейтинг только в 2021 – 2022 годы, из-за этого выборка смещена в сторону именно этих годов. В итоге общее число используемых наблюдений составляет 160. При этом эмитенты чаще получают кредитный рейтинг именно в 4 квартале каждого года. Распределение наблюдений по годам представлено на рисунке 1.

Рис. 1 Распределение наблюдений по годам.

Источник: расчеты автора

В выборку не были включены финансовые компании, потому что подходы к моделированию их кредитных рейтингов значительно отличаются, что исключает возможность их моделирования совместно с нефинансовыми компаниями (Parisa, 2020). Кредитные рейтинги различных рейтинговых агентств сопоставлялись на основе соответствующей методологии Центрального Банка Российской Федерации (Банк России, 2021). При этом из выборки были исключены эмитенты, чей рейтинг выше или равен «A-» по российской рейтинговое шкале. Это объясняется тем, что такие эмитенты имеют высокий уровень кредитоспособности, что противоречит определению высокодоходных облигаций. Согласно методологии Cbonds в индекс Cbonds-CBI RU High Yield (Cbonds, 2022) в некоторых случаях могут включаться эмиссии облигаций, не удовлетворяющих всех критериев высокодоходных облигаций, однако наиболее близкие к ним, с целью расширения базы расчета индекса. Однако ввиду «нетипичности» таких значений их не стоит рассматривать в рамках текущего исследования.

Данные являются квартальными, так как финансовое положение эмитентов может меняться в течение года, поэтому квартальные данные являются наиболее информативными. В случаях, когда эмитент получил кредитный рейтинг в течение квартала, использовался период, предшествующий кварталу получения кредитного рейтинга. Это объясняется тем, что кредитный рейтинг присваивается на основе последней имеющейся финансовой отчетности, что делает некорректным использование будущих значений. Если компания получила кредитный рейтинг от нескольких рейтинговых агентств в течение одного квартала, то использовался наименьший кредитный рейтинг.

Квартальная финансовая отчетность компаний по стандартам РСБУ и их кредитный рейтинг выгружалась из базы данных Cbonds из карточки эмитента.

Помимо показателей, рассчитываемых на основе финансовой отчетности в модель, включены нефинансовые показатели. Данные по ним собирались на основе базы данных СПАРК. Используемые макроэкономические были собраны на основе данных с Росстата.

В целях моделирования кредитного рейтинга эмитентов облигаций были использованы следующие группы факторов: финансовые, нефинансовые, макроэкономические (методологию расчета см. в приложении 2).

Финансовые показатели, используемые для построения классификационной модели, можно условно разделить на 4 подгрупп:

* 1. *Ликвидность активов.* Она оказывает прямое влияние на кредитный рейтинг эмитентов облигаций, определяя ее способность исполнять свои обязательства, так как отражает способность компании в короткие сроки преобразовать имеющиеся активы в денежные средства при необходимости. Чем выше ликвидность активов, тем будет выше кредитный рейтинг эмитента при прочих равных. В рамках данного исследования используются наиболее широко используемые исследователями соотношения, характеризующие ликвидность активов: Cash to Assets, Current Assets to Current Liabilities, NWC to Assets.
  2. *Долговая нагрузка компании.* Данная группа показателей указывает на то, насколько закредитованной является компания, что непосредственно влияет на ее способность исполнять свои обязательства перед займодателями и инвесторами. Были отобраны следующие соотношения, относящиеся к данной группе показателей: Cash to Total Debt, Total Debt to Revenue, Debt to Equity, Debt to EBIT. Также в модель было включено отношение EBIT to Interest, которое характеризует способность компании обслуживать текущие обязательства.
  3. *Рентабельность.* Данная группа факторов отражает способность компании генерировать прибыль относительно её размера. Высокая рентабельность может свидетельствовать о том, что эмитент является финансово устойчивым, что положительно влияет на кредитный рейтинг эмитента. В качестве показателей рентабельности были выбраны: ROA и ROE.
  4. *Прочее.* Помимо этого, в модель были включены показатели, прямо не характеризующие финансовое положение компании и способность исполнять свои финансовые обязательства. Однако они характеризуют бизнес-модель организации, которая может влиять на кредитный рейтинг. В качестве таких показателей были выбраны: NWC to Revenue, EBIT to Revenue, Equity to Asset, Retained Earnings to Assets, EBIT to Assets.

Все вышеуказанные показатели относятся к наиболее часто встречающимся финансовым показателям (Parisa, 2020; Cheng, 2020), однако в модель не были включены показатели, характеризующие денежный поток компании, которые тоже применяются рядом исследователей. Это объясняется тем, что согласно российскому законодательству, отчет о движении денежных средств по РСБУ относится к годовой форме отчетности, и компании не публикуют его на квартальной основе (КонсультантПлюс, б.д.). Поэтому их нельзя будет точно рассчитать для каждого квартала. Помимо этого, в модель не были включены показатели, рассчитываемые на основе характеристик акций эмитента, так как российские эмитенты ВДО, как правило, не имеют обращающихся на бирже акций (Теплова & Родина, 2019).

В модель также были включены наиболее часто встречающиеся в исследованиях макроэкономические показатели: уровень инфляции (для наиболее корректного отражения данного показателя использовался годовой ИПЦ, рассчитываемый по отношению к соответствующему кварталу предыдущего года (Jing-Ping и др., 2020)), годовое изменение реального ВВП по отношению к соответствующему кварталу предыдущего года, уровень безработицы. Первые два показателя рассчитываются в годовом выражении по отношению к кварталу предыдущего года, чтобы избежать влияния сезонности. Использование макроэкономических показателей приводило к повышению точности предсказаний моделей для эмитентов развивающихся стран, к которым относится Российская Федерация (Karminsky и др., 2019; Jing-Ping и др., 2020).

Помимо этого, в модель были включены следующие нефинансовые показатели, которые могут влиять на кредитный рейтинг эмитентов облигаций: отрасль эмитента, наличие значимой исковой нагрузки, отношение компании к малым и средним предприятиям. Эти показатели используются всеми аккредитованными в России кредитными рейтинговыми агентствами при присвоении кредитных рейтингов (НКР, 2023; НРА, 2023; АКРА, 2023; Эксперт Ра, 2023).

## 2.3 Используемые модели предсказания кредитных рейтингов и методы оценки их качества

*Логистическая регрессия (LR)* является статистическим методом, подходящим для решения задач бинарной и мультиклассовой классификации. В основе этого метода лежит логистическая функция, которая описывается следующей формулой:

На её основе логистическая регрессия предсказывает вероятность принадлежности к определенному классу, так сама формула логистической регрессии в общем случае выглядит следующим образом:

Где P – вероятность принадлежности к классу 1*, xn –* значение зависимой переменной *n*-го признака,  *βn* – коэффициент регрессии для n-го признака.

В общем случае логистическая регрессия используется в задачах бинарной классификации, согласно её формуле. Для задач мультиклассовой классификации применяется принцип One-vs-All, при котором для каждого класса строится отдельная логистическая регрессия, предсказывающая вероятность отношения к определённому классу по отношению ко всем остальным классам.

Для повышения обобщающей способности модели к логистической регрессии применяют регуляризацию, которая добавляет к функции потерь логистической регрессии «штраф» за размер весов признаков логистической регрессии. В основном применяют два типа регуляризации: L1 и L2.

Для L1-регуляризации функция потерь логистической регрессии выглядит следующим образом:

Для L2-регуляризации функция потерь логистической регрессии выглядит следующим образом:

где *yi*– значение зависимой переменной для *i*-го наблюдения, σ(zi) – логарифмическая функция, zi – линейная комбинация признаков,βj – коэффициент регрессии для n-го признака, λ – коэффициент регуляризации.

Как видно из вышеуказанных формул разница в L1 и L2 регуляризации заключается в том, что в первом случае суммируются модули коэффициентов регрессии, а во втором – квадраты коэффициентов. Изменение гиперпараметра позволяет регулировать размер «штрафа» в модели.

Так, логистическая регрессия обладает большой обобщающей способностью, однако она чувствительна к выбросам и не способна улавливать нелинейные связи, что ограничивает эффективность применения этого метода на практике.

*Метод опорных векторов (SVM)* является одним из наиболее широкораспространённых методов машинного обучения при решении задач классификации. В его основе лежит поиск гиперплоскости, наиболее точно разделяющей данные на искомые классы.

Алгоритм поиска основан на максимизации расстояний между гиперплоскостью и точками каждого класса, эти расстояния называются опорными векторами. При этом функции гиперплоскости могут быть разными, их называют ядрами. Так, наиболее используемыми являются:

* 1. Линейное. Не изменяет исходное пространство данных, наиболее подходит для линейно разделимых выборок
  2. Полиноминальное. Изменяет исходное пространство на пространство с большей размерностью с помощью полиноминальной функцией, что делает его подходящим для нелинейно разделимых выборок. Полиноминальная функция выглядит следующим образом:

где n – степень многочлена, an – коэффициенты многочлена

* 1. Сигмоидное. Тоже изменяет пространство данных, но с помощью ранее описанной сигмоидной функции.

С точки зрения достоинств метод SVM обладает высокой обобщающей способностью, а также он позволяет строить точные предсказания в тех случаях, когда связи в модели являются нелинейными. Однако этот метод неустойчив к выбросам и масштабу данных, что требует предварительной обработки данных.

*Метод ближайших соседей (KNN)* относится к достаточно простым методам машинного обучения для решения задач классификации. Алгоритм метода устроен следующим образом: класс наблюдения присваивается на основе наиболее часто встречающегося класса ближайших соседей в пространственной плоскости известных наблюдений. Количество ближайших соседей, на основе которых строится предсказание, устанавливается исследователем на основе подбора гиперпараметра, как правило на основе метода сплошного подбора GridSearch.

Данный метод часто демонстрирует высокую предсказательную способность, однако в случаях, когда большое число факторов оказывают влияние на предсказываемую величину, то эффективность метода сильно снижается. При этом метод крайне чувствителен к масштабу данных, что требует предобработки данных.

*Решающее дерево (DT) и Случайный лес (RF).* Данные методы основаны на индукции правил, что достаточно сильно отличает их от методов, описанных ранее. Решающие деревья строятся следующим образом: на основе обучающей выборке алгоритм подбирает правило, на основе которого разобьет выборку на 2 части, а потом будет генерировать правила для того, чтобы разделить каждое полученное подмножества на 2 части, алгоритм продолжается до тех пор, пока не будет достигнут критерий остановки – это может быть глубина дерева (максимальный путь между корнем дерева до листа), минимальное число наблюдений в каждом листе и так далее. Каждому листу соответствует определённый класс. При этом метод решающих деревьев имеет серьезный недостаток: они очень сильно подстраиваются под данные, то есть переобучаются. Для преодоления этой проблемы применяется метод «случайный лес», который включает в себя большое число решающих деревьев, правила которых обобщаются, что часто позволяет снизить переобучение, однако требует большей вычислительной мощности.

Данные методы являются одними из наиболее универсальных, потому что не чувствительны к выбросам, масштабу данных, типу связей между зависимой переменной и независимыми. Однако выработка правил может быть неэффективна на маленьких выборках, а также такие модели склонны к переобучению.

Оценка качества предсказаний моделей в рамках данного исследования осуществляется на основе следующих метрик: Accuracy, ROC-AUC, F1-score.

Метрика Accuracy (точность) под собой подразумевает долю верных ответов модели по отношению к неправильным ответам. Несмотря на то, что данная метрика является интуитивно понятной, при сильно несбалансированных выборках данная метрика плохо отражает реальное качество модели, что вносит ограничения в ее применимость. Модель обладает высокой предсказательной способностью в рамках задачи моделирования кредитного рейтинга, если превышает порог в 55% (Parisa, 2020).

Метрика ROC-AUC рассчитывается как площадь под графиком ROC-curve, который отражает зависимость True Positive Rate (TPR) от False Positive Rate (FPR) при изменении порога принятия решения. TPR и FPR — это доли верно и ложно классифицированных соответственно положительных объектов. Наибольшее качество модели соответствует значению ROC-AUC в 1, а наименьшее в 0,5. По умолчанию рассчитывается для бинарного случая, однако применяется и в задачах мультиклассовой классификации путем усреднения следующими способами: микроусреднение (при расчете итогового значения метрики составляющие части суммируются для каждого класса и на основе этого рассчитывается итоговая метрика), макроусреднение (итоговые значения метрики усредняется для каждого класса с одинаковыми весами), взвешенное (итоговые значения метрики усредняется для каждого класса с весами, соответствующими доле класса в выборке).

Метрика F1-score является одной из наиболее универсальных метрик оценки качества моделей, так как позволяет учитывать и «чувствительность» модели, и «полноту» предсказания. Рассчитывается данная метрика на основе матрицы ошибок по следующей формуле:

где *TP* – количество верно определенных положительных результатов, *FP и FN* – количество неверно определенных положительных и отрицательных результатов соответственно.

По аналогии с ROC-AUC, F1-score в общем случае рассчитывается для задач бинарной классификации. Для задач мультиклассовой классификации применяются методы усреднения, описанные ранее.

## 2.5 Дескриптивный и корреляционный анализ

Все последующие расчеты проводились на основе языка программирования Python 3.9 в программе Jupyter Notebook. Перед построением моделей был проведен дескриптивный анализ, основные статистики переменных в выборке представлены в таблице 1.

Таблица 1. Дескриптивные статистики переменных.



Источник: расчеты автора

Согласно таблице 1 среди макроэкономических и нефинансовых переменных выбросов и аномалий нет. Стоит отметить, что 40% наблюдений принадлежит эмитентам, имеющих значимую исковую нагрузку. И 50% наблюдений относится к эмитентам, являющимися МСП (малыми и средними предприятиями). Это говорит о том, что вышеуказанные дамми-переменные являются достаточно вариативными и поэтому их включение в последующую модель является целесообразным.

Также из дескриптивных статистик, указанных в таблице 1, можно сделать вывод, что в данных очевидно присутствуют выбросы во всех финансовых переменных. Это говорит о необходимости проведения обработки выбросов, так как большая часть используемых методов неустойчивы к выбросам. В приложении 3 представлены графики распределения переменных до обработки выбросов. Они подтверждают тот факт, что в данных присутствуют выбросы во всех финансовых показателях.

В рамках данной задачи удаление выбросов является нецелесообразным, так как в выборке присутствует большое их число, удаление которых приведет к сокращению имеющейся выборки, а также может и к её искажению. Поэтому выбросы были заменены медианами соответствующих показателей, такой подход позволит минимизировать влияние выбросов на итоговую модель. Для определения выбросов использовался межквартильный критерий, согласно которому выбросом признается наблюдение, которое отличается в большую/меньшую сторону от 1 и 3 квартиля распределения соответственно более чем в полтора межквартильных расстояний (расстояние между 1 и 3 квартилями распределениями). В приложении 4 представлены графики распределения финансовых показателей после обработки выбросов. После удаления выбросов аномалий в распределениях обнаружено не было и распределение стали похожими на нормальное.

Помимо этого, в переменных заметно, что средние и медианы колеблются относительно друг друга, а так как некоторые используемые модели чувствительны к масштабу данных, следует провести масштабирования данных методом Scaling, которое позволит привести данные к единому масштабу без изменения характера распределения.

Нефинансовая переменная SECTOR содержит текстовые значения, поэтому для её использования в математических моделях необходимо привести переменную к числовому формату. Однако выбор метода кодирование текстовых признаков зависит от характера распределения классов. Распределение наблюдений выборки по отраслям эмитентов представлено на рисунке 5.

Рис. 5 Распределение наблюдений по отраслям эмитентов. Источник: расчеты автора.

Согласно рисунку 5 около половины наблюдений приходится на 4 следующие отрасли: «Строительство зданий», «Оптовая торговля», «Прочее машиностроение и приборостроение», «специализированная розничная торговля». Помимо того, что классы распределены крайне неравномерно, почти все классы представлены малым числом значений (меньше 5). Кодирование показателя отрасли было произведено с помощью метода OneHotEncoder, создающего дамми-переменные для каждой категории за исключением одной во избежание возможности точного выражения каждой переменной из другой, приводящего к не единственности решения.

Зависимая переменная «кредитный рейтинг» изменяется в пределах от 4 до 7. При этом стоит отметить, что, так как используется рейтинговая шкала Центрального Банка Российской Федерации для сопоставления кредитных рейтингов, присваиваемых различными российскими рейтинговыми агентствами, то минимальное значение будет соответствовать наивысшему значению кредитного рейтинга (Банк России, 2021). Среднее значение кредитного рейтинга в 5.46, что не указывает на явное смещения выборки в сторону эмитентов с более высоким или низким качеством, однако, данный факт, требует дополнительной проверки. Распределение кредитных рейтингов представлены на рисунке 4.

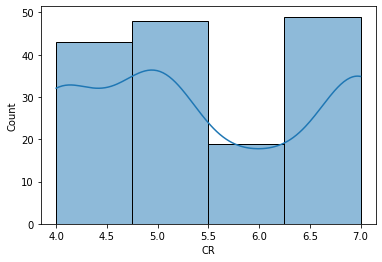


Рис. 4 Распределение кредитного рейтинга. Источник: расчеты автора.

Из рисунка 4 видно, что классы, распределены не равномерно, что говорит о необходимости использования взвешенного или макроусреднения при оценке последующих моделей. Однако стоит отметить, что встречаемость только одного класса значительно отличается от остальных. Это может привести к тому, что при моделировании этот класс будет предсказываться хуже остальных, так как он встречается значительно реже остальных. Однако данное исследование не сфокусировано именно на предсказании «редкого» класса, а на построении наиболее универсальной модели, поэтому является целесообразным использование взвешенного усреднение метрик качества моделей, чтобы не переобучать модель на «редком» классе. Но также стоит сопоставлять взвешенное F1 c макроусредненным F1, чтобы не допустить ситуации игнорирования редкого класса.

Таким образом, на основе дескриптивного анализа была проведена обработка данных, которой должно было достаточно для построения моделей, предсказывающих кредитный рейтинг эмитентов. Помимо этого, была выявлена неравномерность распределения зависимой переменной, что обусловило использование взвешенного усреднения метрик качества моделей. Еще одним важным выводом дескриптивного анализа является тот факт, что переменные, характеризующие исковую нагрузку компаний и их наличие статуса МСП является достаточно вариативными, поэтому потенциально могут повысить предсказательную способность моделей, однако переменная, отражающая отрасль компаний представлена в основном несколькими классами, что может значительно ограничить применимость переменной SECTOR.

Далее был проведён корреляционный анализ. Тепловая корреляционная карта (за исключением переменных сектора) представлена на рисунке 5. С точными значениями корреляции можно ознакомиться в приложении 5.

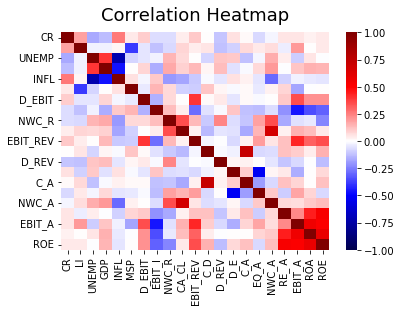


Рис.5 Корреляционная тепловая карта. Источник: расчеты автора

Значение силы корреляции зависимой переменной с каждой независимой переменной является слабым и не превышает 0,26. Наиболее коррелирующими переменными являются макроэкономические (сила корреляции (модуль корреляции) составляет от 0,17 до 0,26) и переменная, характеризующая наличие исковой нагрузки (сила корреляции составляет 0,19). Сила корреляции финансовых показателей находится на очень низком уровне и колеблется от 0,03 до 0,11. Наиболее коррелирующими финансовыми переменными являются: «EBIT\_REV», «D\_REV», «EQ\_A» (значение корреляции составляет 0,11, -0,11, -0,7 соответственно). Учитывая тот факт, что все эти переменные используются кредитными рейтинговыми агентствами при присвоении кредитных рейтингов, стоит ожидать, что низкая сила корреляции связана с тем, могут иметь место нелинейные связи между факторами и кредитным рейтингом, что как раз может приводить к низкому уровню корреляции. Также низкий уровень корреляции может быть вызван тем, что каждый фактор по отдельности оказывает слабое влияние на итоговый кредитный рейтинг и необходимо рассматривать все показатели в совокупности, что соответствует тому, как кредитные рейтинговые агентства присваивают кредитные рейтинги.

Из рисунка 5 также видно в переменных присутствуют переменные, сильно коррелирующие между собой. Такие переменные необходимо удалить во избежание мультиколлинеарности. Поэтому из выборки были исключены переменные, обладающие средним и высоким уровнями корреляции (больше 0.5) с другими независимыми переменными. А именно: «UNEMP», «GDP», «NWC\_A», «C\_A», «D\_E», «RE\_ A», «EBIT\_A», «ROA». Из каждой группы сильно коррелирующих между собой переменных были исключены те, которые в меньшей степени коррелировали с кредитным рейтингом и/или коррелировали с большим числом других независимых переменных. Высокий уровень корреляции исключённых финансовых переменных обусловлен использованием одних и тех же статей отчетности при расчете показателей и относительно малой вариативностью некоторых статей, а высокий уровень корреляции макроэкономических показателей обусловлен общеизвестной взаимосвязи этих показателей.

Корреляции переменных отраслей с кредитным рейтингом представлены в приложении 6. Среди наиболее встречающихся отраслей уровни корреляции представлены в таблице 2.

Таблица 2. Значения корреляции между переменными принадлежности к определенной отрасли и переменной кредитного рейтинга.

|  |  |
| --- | --- |
| **Отрасль** | **Корреляции** |
| Строительство зданий | -0,30 |
| Оптовая торговля | 0,28 |
| Прочее машиностроение и приборостроение | 0,17 |
| Специализированная розничная торговля | -0,15 |
| ИТ сервисы и программное обеспечение | -0,08 |
| АПК и сельское хозяйство | 0,17 |
| Производство строительных материалов | -0,16 |
| Производство продуктов и напитков | 0,23 |
| Производство лекарств и биотехнологии | -0,10 |
| Образование | 0,04 |

Источник: расчеты автора.

Из таблицы 2 видно, что принадлежность к отраслям «строительство зданий», «производство строительных материалов», «специализированная розничная торговля» связана с более высоким уровнем кредитного рейтинга. При этом принадлежность к «оптовой торговле», «производству продуктов и напитков», «прочему машиностроению и приборостроению» и «АПК и сельскому хозяйству» связано с менее высоким кредитным рейтингом. Уровни корреляции между этими переменными являются слабыми, однако в основном выше, чем по остальным переменным, что говорит о необходимости включения этих переменных в модель, что соответствуют существующим исследованиям (Perry, Henderson, Jr, & Cronan, 1984). Переменные отраслей, не включенных в таблицу 2, исключались, так как к этим отраслям относилось менее 4 наблюдений, то есть имела место низкая вариативность признака, что не позволит ей повысить качество предсказания, но создаст лишние переменные, что может негативно сказаться на качестве модели.

Помимо этого, для учитывания временного эффекта использовалась закодированная по принципу метода TargetEncoder переменная, характеризующая квартал получения кредитного рейтинга. Это обусловлено тем, что с 2018 по 2022 на российском финансовом рынке произошел ряд «потрясений», что обуславливает необходимость отражение временного эффекта.

## 2.7 Результаты моделирования кредитных рейтингов

### 2.7.1 Подбор гиперпараметров

Перед описанием результатов моделей стоит описать, какие гиперпараметры моделей были подобраны методом GridSearchCV. В некоторых случаях подобранные гиперпараметры могут характеризовать особенности связей в модели.

*Логистическая регрессия* является достаточно универсальным статистическим методом, поэтому малое число гиперпараметров может быть подобрано: метод регуляризации (L1 или L2), а также λ, которая характеризует размер «штрафа», накладываемого регуляризацией. Так, наилучшей оказалась регуляризация L1 c λ равной 1,5. Особенность такой регуляризации заключается в том, что у малозначимых факторов модели коэффициенты зануляются, однако в рамках данной задачи удалять переменные не стоит, потому что при мультиклассовой классификации строится количество логистических регрессий, равное количеству предсказываемых классов. Так при предсказании разных классов отдельные переменные могут оказаться значимыми. Помимо этого, при использовании других методов моделирования новые переменные могут оказаться значимыми для предсказания, поэтому желательно не исключать объясняющие переменные на основе результатов логистической регрессии.

*Метод опорных векторов.* Были подобраны гиперпараметры для 2 вариаций данного метода: с линейным и нелинейными ядрами. Наибольшее качество получила модель с линейным ядром с параметром «ширина разделяющей прямой» равным 1,6. Наилучший метод опорных векторов с нелинейным ядром использует полиноминальное ядро и «ширину разделяющей прямой» равной 18.

*Метод ближайших соседей.* Данный метод очень чувствителен к подбору гиперпараметра, характеризующего количество «ближайших соседей». Так, наилучшее качество показала модель со значением данного гиперпараметра равное 7. Это хорошее значение, так как в каждую группу в идеале должны попадать разные компании, а значение равное 7 подразумевает, что в каждой группе будет по несколько компаний.

*Решающее дерево.* Данный метод аналогично предыдущему зависим от подобранных гиперпараметров. Главным к подбору является гиперпараметр, характеризующий минимальное необходимое число наблюдений в каждом «листе» дерева. Наилучшая модель имеет количество наблюдений в каждом «листе» равное 10. Это хорошее значение, так как в каждый «лист» попадают различные компании.

*Случайный лес.* Является обобщением предыдущего метода, поэтому подбирается тот же гиперпараметр и само количество деревьев. Минимальное количество наблюдений в каждом листе составило 9, а количество деревьев в «лесу» 20, что также как и в предыдущем случае является хорошим значением.

Ввиду большого числа признаков относительно размера выборки был проведен жадный отбор признаков методом recursive feature elimination (RFE). Однако исключение переменных не привело к повышению качества моделей, а только к его снижению, тем самым, было решено более не исключать переменные.

### 2.7.2 Оценка качества предсказаний моделей

Основные метрики качества полученных моделей представлены в таблице 3. Оценка качества моделей оценивалась по всей выборке, однако обучение проводилось только на 60% выборки, что позволяет избежать переобучения модели, чтобы адекватно оценить качество модели. Из нее видно, что наилучшими моделями оказались логит-модель и SVM, включающие финансовые, нефинансовые и макроэкономические показатели. Логистическая регрессия несмотря на то является достаточно простым методом смогла продемонстрировать высокое качество предсказания: точность составила 64%, взвешенное F1 – 0,63, а ROC AUC – 0,82. Учитывая тот факт, что задача сведена к предсказанию 4 классов, то качество модели является достаточно высоким. Таким образом, можно утверждать, что гипотеза 1 не отвергается.

Таблица 3. Основные метрики качества предсказаний моделей.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Модель | Метрики качества | | | |
| Accuracy | ROC AUC | F1\_weighted | F1\_macro |
| Log\_All | 0,64 | 0,82 | 0,63 | 0,61 |
| Log\_Fin | 0,4 | 0,61 | 0,4 | 0,38 |
| SVM\_Lin\_ALL | 0,65 | 0,83 | 0,64 | 0,62 |
| SVM\_Lin\_Fin | 0,35 | 0,57 | 0,35 | 0,34 |
| SVM\_Poly\_ALL | 0,59 | 0,7 | 0,59 | 0,57 |
| SVM\_Poly\_Fin | 0,34 | 0,53 | 0,29 | 0,29 |
| KNN\_All | 0,55 | 0,77 | 0,53 | 0,48 |
| KNN\_Fin | 0,4 | 0,67 | 0,36 | 0,32 |
| DT\_All | 0,57 | 0,76 | 0,54 | 0,46 |
| DT\_Fin | 0,47 | 0,65 | 0,41 | 0,35 |
| RF\_All | 0,55 | 0,81 | 0,51 | 0,46 |
| RF\_Fin | 0,45 | 0,75 | 0,35 | 0,29 |

Источник: расчеты автора

При этом использование исключительно финансовых переменных значительно снижает предсказательную способность всех моделей, что соответствует результатам последних исследований (Jing-Ping и др., 2020), а это позволяет утверждать, что гипотеза 4 не отвергается.

С точки зрения методов машинного обучения наилучшим оказался метод опорных векторов с линейным ядром, использующим все рассматриваемые переменные, однако его качество практически идентично логит-модели. Это является нетипичным результатом для большинства исследований, однако Wallis и др. (2019) и Ye и др. (2008) получили схожие результаты для выборок американских компаний. То есть гипотеза 2 отвергается, методы машинного обучения не привели к росту предсказательной способности модели.

Применение нелинейных ядер не позволило повысить качество модели, а только снизило его. При этом и другие методы, способные выявлять нелинейные связи не смогли получить высокое качество предсказания. Так, методы KNN, RF и DT продемонстрировали качество меньшее, чем логит-модель и SVM с линейным ядром. При этом для методов KNN, RF и DT справедливо суждение, что взвешенный F1 больше, чем макроусредненный F1, это вероятно вызвано тем, что эти методы хуже остальных предсказывают кредитный рейтинг «6», который встречается реже остальных, в то время как линейные методы лучше справляются с предсказанием этого класса. Невысокое качество предсказания решающего дерева и соответственно случайного леса является достаточно частым результатом. Так (Shin & Han, 2001) столкнулись с аналогичным результатом на выборке Корейских компаний, а (Wallis и др., 2019) на выборке американских компаний. Таким образом, использование нелинейных методов предсказания кредитного рейтинга эмитентов не позволяет повысить качество предсказаний модели в рамках данного исследования, то есть гипотеза 3 отвергается.

Для лучшей интерпретации результатов предсказаний кредитного рейтинга стоит рассмотреть по классовое качество предсказания, а также матрицу ошибок. Рассмотреть стоит именно логит-модель, потому что она имеет высокое качество выше, чем по остальным моделям (за исключением SVM, где качество предсказания практически идентично). При этом логит-модель является более интерпретируемой и простой для построения. В таблице 3 представлена расширенная статистика по метрике F1 логит-модели, использующей финансовые, нефинансовые, макроэкономические объясняющие переменные, на основе всех наблюдений.

Таблица 4. Поклассовая оценка качества логит-модели.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Class** | **Precision** | **Recall** | **F1-score** | **Support** |
| **4** | 0,69 | 0,77 | 0,73 | 43 |
| **5** | 0,68 | 0,4 | 0,5 | 48 |
| **6** | 0,46 | 0,58 | 0,51 | 19 |
| **7** | 0,64 | 0,78 | 0,7 | 49 |
|  | | | | |
| **Accuracy** |  |  | 0,64 | 159 |
| **Macro avg** | 0,62 | 0,63 | 0,61 | 159 |
| **Weighted avg** | 0,64 | 0,64 | 0,63 | 159 |

Источник: расчеты автора

Из таблицы 4 видно, что логит-модель лучше всего предсказывает «крайние значения», то есть кредитные рейтинги «4» или «7». Этот факт может быть объяснен тем, что на концах рейтинговой шкалы находятся наиболее легкие к предсказанию наблюдения, так как кредитное качество эмитентов с кредитными рейтингами «4» и «7» могут сильнее отличаться от соседних кредитных рейтингов, чем классы «5» и «6» между собой. При этом кредитный рейтинг «6» встречается реже остальных (более чем в 2 раза относительно каждого из остальных классов), что также снижает предсказательную способность этого кредитного рейтинга, что и отразилось в метриках качества данного кредитного рейтинга. Стоит отметить, что для кредитных рейтингов «4» и «7» метрики «precision» и «recall» отличаются между собой не очень сильно, что свидетельствует о достаточно хорошем балансе между чувствительностью и полнотой. Для классов «5» и «6» ситуация иная: «5» класс имеет излишне высокую чувствительность относительно полноты, то есть модель склонна распределять истинные «5» по другим классам, а «5» класс присваивать только в случаях наибольшей уверенности модели. В случае с «6» классом ситуация обратная: модель слишком склонна присваивать «6» даже тем наблюдения, которые не относятся к данному классу. Ввиду того, что рассматриваемые классы являются рейтингами и могут быть ранжированы, то стоит рассмотреть матрицу ошибок, представленную на рисунке 6, так как ошибка завышения и занижения кредитного рейтинга не являются равноценными. Для инвесторов и эмитентов желательно получить заниженный рейтинг, чем завышенный, так как завышенный рейтинг может привести к излишнему оптимизму и недооценке рисков, который может привести к значительным финансовым потерям.



Рис.6 Матрица ошибок логит-модели. Источник: расчеты автора

Из рисунка 6 видно, что истинный «4» кредитный рейтинг классифицируется моделью достаточно корректно, небольшое число наблюдений, относящиеся к «4» кредитному рейтингу, были классифицированы неверно, были отнесены к остальным рейтингам. Хотя модель также часто присваивала «4» кредитный рейтинг наблюдениям, которые на самом деле относятся к «5» классу. Истинный «5» и «6» кредитный рейтинг, как было описано ранее предсказывается моделью хуже. Те наблюдения, которые модель отнесла к «5» кредитному рейтингу были отнесены достаточно точно, однако большое число наблюдений истинно относящихся к «5» кредитному рейтингу были отнесены к другим рейтингам, особенно часто к «4». Истинный «6» кредитный рейтинг предсказывается моделью более точно, большая часть истинных «6» кредитных рейтингов была классифицирована верно по сравнению с «5» кредитным рейтингом. Истинный «7» кредитный предсказывается несколько хуже остальных. Небольшое число наблюдений истинно относящихся к «7» кредитному рейтингу равномерно распределены по остальным классам.

Таким образом, на основе анализа корреляционной матрицы можно сделать вывод, что построенная модель в случае ошибок в основном предсказывает наиболее близкие к истинному значения, что подтверждает применимость линейных методов для моделирования кредитных рейтингов.

На рисунке 7 представлена ROC кривая. Ее анализ позволит более точно оценить общее качество предсказания логит-модели, не учитывая направление отклонения в неверно классифицированных наблюдениях, но учитывая уверенность алгоритма.

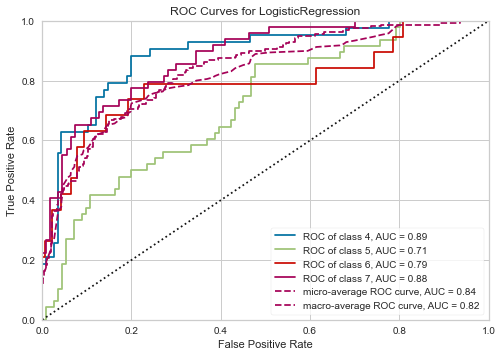


Рис.7 ROC-кривые логит-модели. Источник: расчеты автора

Из рисунка 7 видно, что все сделанные выводы являются релевантными. Кредитный рейтинг «4» и «7» предсказываются моделью наиболее точно. Кредитный рейтинг «6» предсказывается несколько хуже «крайних» рейтингов, однако близок к ним. Кредитный рейтинг «5» предсказывается достаточно хорошо, однако значительно хуже, чем по остальным кредитным рейтингам, что подтверждает ранее описанные выводы.

Таким образом, можно сделать вывод, что моделирование кредитного рейтинга российских эмитентов высокодоходных облигаций возможно с помощью математических методов, при этом предсказательная способность таких моделей будет достаточно высокой. Однако стоит учитывать, что использование исключительно финансовых показателей значительно сокращает предсказательную способность моделей. При этом на примере выборки российских эмитентов облигаций является достаточным использование логистической регрессии, в то время как применение методов машинного обучения в том числе выявляющих нелинейных связей не позволяет значительно повысить качество предсказания.

## 2.8 Выводы

Актуальность данной работы обусловлена тем фактом, что все больше российских частных инвесторов выходит на финансовый рынок, и при этом они в основном полагаются на стороннюю аналитику и различные кредитные рейтинги. Кредитные рейтинги являются одним из наиболее авторитетных источником информации, на которые опираются как частные инвесторы, так и государственные регуляторы. Однако дороговизна получения кредитного рейтинга приводит к тому, что эмитенты не могут регулярно обновлять свои кредитные рейтинги, из-за чего они могут быстро устаревать, эта проблема особенно остро стоит для эмитентов высокодоходных облигаций. При этом высокодоходные облигации являются инструментом, который все больше привлекает внимание частных инвесторов. Эти факты обуславливают исследовательский и практический интерес к моделированию кредитных рейтингов российских эмитентов высокодоходных облигаций. В рамках данного исследования была рассмотрена история рынка высокодоходных облигаций и его особенностей, обусловивших интерес инвесторов к нему, а таже роль кредитных рейтинговых агентств в его развитии. Так, было выявлено, что при предсказании кредитных рейтингов эмитентов ВДО необходимо использовать широкий спектр переменных: финансовые, нефинансовые и макроэкономические показатели, определяющие кредитный рейтинг. Для проведения исследования была собрана выборка по российским эмитентам ВДО, имеющим кредитный рейтинг и относящимся к нефинансовым отраслям, за период с 2018 по 2022 годы.

Несмотря на то, что данное исследование использует наиболее универсальные подходы, продемонстрировавшие хорошую применимость на различных выборках, оно имеет ряд ограничений, которые могут вызваны объективными причинами, а также могут быть проработаны в будущих исследованиях. Так, можно выделить следующие группы ограничений.

*Небольшая выборка.* Данное исследование сфокусировано на российских эмитентах ВДО в то время, как российский рынок ВДО представлен достаточно небольшим числом эмитентов. Это приводит к тому, что часть отраслей представлены малым числом компаний, что ограничивает предсказательную способность моделей. Данная проблема может быть преодолена путем поиска выборок иностранных эмитентов ВДО, чьи компании наиболее близки к российским. Однако данная задача выходит за рамки этого исследования и требует иного подхода.

*Невозможность точного сопоставления рейтинговых шкал и недостаток исторических данных.* В рамках данного исследования используется шкала для сопоставления кредитных рейтингов разных рейтинговых агентств, разработанная Центральным Банком Российской Федерации (Банк России, 2021). Она объединяет по 1-2 уровня кредитного рейтинга в одну «ступень». Это подход упрощает задачу моделирования, однако не позволяет отображать единичные отклонения в кредитном рейтинге. Однако на данный момент времени использование более точного сопоставления кредитных рейтингов разных КРА невозможно ввиду отсутствия исторических данных дефолтов эмитентов ВДО в России.

*Необходимость дальнейшего расширения используемых показателей.* В рамках данного исследования использовались наиболее часто используемые объясняющие переменные. Учитывая тот факт, что результаты данного исследования не полностью согласуются с исследованиями, посвященные другим финансовым рынкам, релевантной задачей будет выявление специфичных для российского рынка переменных, способных влиять на кредитный рейтинг. Помимо этого, также следует рассмотреть переменные, лучше характеризующие бизнес-модель каждого из российских эмитентов.

*Низкая интерпретация методов машинного обучения.* Несмотря на то, что методы машинного обучения потенциально обладают высокой предсказательной силой, на их основе затруднительно делать какие-либо выводы по поводу значимости тех или иных показателей, что ограничивает применимость результатов на практике.

*Возможное применение более сложных методов.* Для моделирования кредитного рейтинга российских эмитентов ВДО использовались наиболее универсальные и хорошо зарекомендовавшие себя методы, однако возможным направлением развития данного исследования может служить применение более сложных методов глубинного обучения, а также сложных ансамблевых моделей, использование которых может привести к повышению предсказательной способности моделей.

Учитывая вышеизложенные ограничения, в рамках данного исследования были построены различные модели, способные предсказывать кредитный рейтинг российских эмитентов высокодоходных облигаций и сопоставлены качества их предсказаний. Так, использовались следующие методы: логистическая регрессия, метод опорных векторов (с линейными и нелинейными ядрами), метод ближайших соседей, решающее дерево и случайный лес. Оценка качества моделей производилась с использованием различных метрик: Accuracy, F1-score, ROC-AUC; а также на основе анализа матрицы ошибок. Качество предсказания линейных моделей (логит-модель и SVM с линейным ядром) находится на достаточно высоком уровне (точность предсказания составляет около 65%), и поэтому эти модели могут использоваться инвесторами и эмитентами для решения практических задач. Однако точность моделей не близка к 100%, что делает некорректным использование данных моделей как единственного инструмента оценки финансового состояния компании, однако будет полезным их использование в качестве дополнительного инструмента, применяемого на ежеквартальной основе, способного «подсветить» значимые изменения в финансовом состоянии компании.

Помимо практического применения результаты данного исследования могут также быть полезными и для исследователей рынка облигаций в России, в частности высокодоходных облигаций. Так, например, было выявлено что при моделировании кредитного рейтинга российских эмитентов облигаций необходимо помимо финансовых показателей использовать макроэкономические и нефинансовые показатели, в то время как исследования, посвященные финансовым рынкам развитых стран, могли не использовать их и получать модели с высоким качеством предсказания. В случае с российскими эмитентами ВДО использование исключительно финансовых переменных приводит к получению моделей с очень слабой предсказательной силой. Также было продемонстрировано, что линейные методы моделирования кредитного рейтинга российских эмитентов ВДО демонстрируют большее качество по сравнению с нелинейными. Это говорит о том, что российский рынок высокодоходных облигаций имеет свои особенности, которые могут быть рассмотрены в будущих исследованиях.

# Список литературы

1. Карминский, А. М. (2015). *Кредитные рейтинги и их моделирование.* Москва: Издательский дом Высшей школы экономики.
2. Семерина, Ю. В., & Байбеков, И. Р. (2014). Проблемы оценки кредитного качества эмитентов облигаций. *Финансы и кредит, №28 (604)*.
3. Теплова, В. А., & Родина, Т. В. (2019). *Высокодоходные облигации: от истории становления рынка в США до российских реалий.* Москва: ИНФРА-М.
4. Altman, E. I., & Simon, B. E. (1994). The Investment Performance of Defaulted Bonds for 1994 and 1987-1994. *NYU Working Paper No. FIN-94- 004.* [*https://ssrn.com/abstract=1298807*](https://ssrn.com/abstract=1298807).
5. Cheng, F. L. (2020). *Credit Analysis, Bond Rating Forecasting, and Default Probability Estimation.*
6. Daniel, C., Jana, H., Hosn, e. W., Lorena, C., & López, d. R. (2019). Corporate rating forecasting using Artificial Intelligence statistical techniques. *Investment Management and Financial Innovations, 16*, 295-312.
7. Fisher, L. (1959). Determinants of risk premiums on cor-changes. *Journal of Political Economy*, 217-237.
8. Goldstein, I., & Chong, H. (2020). Credit Rating Inflation and Firms' Investments. *The journal of Finance*, 2929-2972.
9. Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies, 33 №. 5.*, 2223-2273.
10. Ha ́jek, P., & Olej, V. ́. (2011). Credit rating modelling by kernel-based approaches with supervised and semi-supervised learning. *Neural Comput & Applic*, 761-773.
11. Hickman, B. (1958). Corporate Bond Quality and Investor Experience. Chapter 3: Agency Ratings. Princeton: Princeton University Press.
12. Horrigan, J. (1996). The determination of long term cre-produce a better prediction model. *Journal of Accounting Research*, 44-62.
13. Hsu-Che, W., Ya-Han, H., & Yen-Hao, H. (2014). Two-stage credit rating prediction using machine learning techniques. *Kybernetes*.
14. Jiexian, Y., Guolan, l., & Hogjin, W. (2021). Credit Grade Prediction Based on Decision Tree Model. *2021 IEEE International Conference on Intelligent Systems and Knowledge Engineering*.
15. Jing-Ping, L., Nawazish, M., Birjees, R., & Deping, X. (2020). Machine learning and credit ratings prediction in the age of fourth industrial revolution. *Technological Forecasting & Social Change*.
16. Kürüm, E., Yildirak, K., & Weber, G.-W. (2011). A classification problem of credit risk rating investigated and solved by optimisation of the ROC curve. *Central European Journal of Operations Research*, 529-527.
17. Karminsky, A. M., Grishunin, S., Dyachkova, N., & Bisenov, M. (2019). The comparison of empirical methods for modeling credit ratings of industrial companies from BRICS countries. Год: 2019. *Eurasian Economic Review A Journal in Applied Macroeconomics and Finance*.
18. Karminsky, A. M., Khromova, E. P., & Kudrov, R. A. (2021). Empirical Modeling of International Banks’ Credit Risk: Assessment and Comparison of Credit Ratings. *Eurasian Business and Economics Perspectives*, 139-161.
19. Khemakhem, S., & Boujelbene, Y. (2015). Credit risk prediction: A comparative study between discriminant analysis and the neural network approach. *Accounting and Management Information Systems, 14(1), 60.*
20. Kräussl, R. (2003). Do Credit Rating Agencies Add to the Dynamics of Emerging Market Crises? *Center for Financial Studies, Working Paper No 2003/18.*
21. Maher, J. J., & Sen, T. K. (1997). Predicting Bond Ratings Using Neural Networks: A Comparison with Logistic Regression. *Intelligent Systems in Accounting, Finance and Management*, 59-72.
22. Mirko, M., Simone, N., Fabio, P., & Gianluca, V. (2019). *Corporate default forecasting with machine learning.* Bank of Italy, Economic Research and International Relations Area.
23. Parisa, G. I. (2020). A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees. *North American Journal of Economics and Finance*.
24. Perry, L., Henderson, G., Jr, & Cronan, T. (1984). Multivariate analysis of corporate bond ratings. *Journal of Intelligent Systems in and industry classifications*, 27-36.
25. Pinches, G., & Mingo, K. (1975). The role of subordination and industrial bond ratings. *Journal of Finance, 28(1)*, 1-18.
26. Pogue, T., & Soldofsky, R. M. (1996). Journal of Finacial and Quantitative Analysis. *‘What’s in a bond rating?*, 201-228.
27. Rosario, M. (2015). High Yield Bond Market: Features and Risks of a Growing Market. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3400677>.
28. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 206-215.
29. Shin, K., & & Han, I. (2001). A case-based approach using inductive indexing for corporate bond rating. *Decision Support Systems, 32(1)*, 41-52.
30. Vazza, D., Kraemer, N., & Gunter, E. (2019). *Annual Global Corporate Default And Rating Transition Study. S&P Global Ratings.*
31. Wallis, M., Kumar, K., & Gepp, A. (2019). Credit rating forecasting using machine learning techniques. Managerial perspectives on intelligent big data analytics. *IGI Global*, 180-198.
32. Wang, G. M., Chen, J. G., & Yang, Y. (2020). Financial distress prediction: Regularized sparse-based random subspace with ER aggregation rule incorporat- ing textual disclosures. *Applied Soft Computing*.
33. Wei, L. (2022). Optimization and Application of Random Forest Algorithm for Applied Mathematics Specialty. *Machine Learning for Security and Communication Networks*.
34. Ye, Y., Liu, S., & Li, J. (2008). A multiclass machine learning approach to credit rating prediction. *IEEE*, 57-61.
35. Zhong, H., Miao, C., Shen, Z., & Feng, Y. (2014). Comparing the learning effectiveness of bp, elm, i-elm, and svm for corporate credit ratings. *Neurocomputing*, 285–295.
36. АКРА. (2023). *Рейтинговые шкалы*. URL: <https://www.acra-ratings.ru/about-ratings/scales/> (Дата обращения: 09.05.2023)
37. Банк России. (2021). *Информация о сопоставлении рейтинговых шкал российских кредитных рейтинговых агентств*. URL: <https://cbr.ru/press/pr/?file=30122021_101000PR2021-12-30T10_03_38.htm&fbclid=IwAR3fIOYibtZ_zZfmXCWSrEW8OQ_kXAJ_6kPwvrTM5oquOSk50ehYrfikGHY> (Дата обращения: 09.05.2023 )
38. Банк России. (2023). *Кредитные рейтинговые агентства*. Кредитные рейтинговые агентства URL: https://www.cbr.ru/finm\_infrastructure/ra/ (Дата обращения: 09.05.2023)
39. КонсультантПлюс. (б.д.). *Справочная информация: "Формы бухгалтерской отчетности"*. Отчет о движении денежных средств (ОКУД 0710005) URL:<http://www.consultant.ru/document/cons_doc_LAW_32453/fbb01d3f17ffb46020bb29eea4d20cb0d0a84c5c/> (Дата обращения: 09.05.2023)
40. Московская Биржа. (б.д.). *Регистрация выпуска биржевых облигаций*. URL: <https://www.moex.com/s1880> (Дата обращения: 09.05.2023)
41. НКР. (2023). Действующие методологии URL: <https://ratings.ru/methodologies/current/> (Дата обращения: 09.05.2023)
42. НРА. (2023). Рейтинговые методологии URL: <https://www.ra-national.ru/methodology/> (Дата обращения: 09.05.2023)
43. Эксперт Ра. (2023). Действующие методологии URL: <https://www.raexpert.ru/ratings/methods/current> (Дата обращения: 09.05.2023)
44. Investopedia. (2022). *High-Yield Bond: Definition, Types, and How to Invest*. URL: <https://www.investopedia.com/terms/h/high_yield_bond.asp> (Дата обращения: 09.05.2023)
45. NASDAQ. (2018). Speculative-grade bond URL: <https://www.nasdaq.com/glossary/s/speculative-grade-bond> (Дата обращения: 09.05.2023)
46. SEC. (2013). What Are High-yield Corporate Bonds? URL: <https://www.sec.gov/files/ib_high-yield.pdf> (Дата обращения: 09.05.2023)

# Приложение

*Приложение 1*

*Таблица 1. Информация о сопоставлении рейтинговых шкал российских кредитных рейтинговых агентств*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Количество ступеней вниз | Описание исходя из рейтинговых категорий рейтинговых шкал | Национальная шкала для Российской Федерации | | | |
| АКРА (АО) | АО «Эксперт РА» | ООО «НКР» | ООО «НРА» |
| 0 | Наивысший / максимальный уровень кредитоспособности / финансовой надежности / финансовой устойчивости | AAA (RU) | ruAAA | AAA.ru | AAA|ru| |
| 1 | Высокий уровень кредитоспособности / финансовой надежности / финансовой устойчивости (от умеренно высокого до высокого) | AA+(RU), AA(RU) | ruAA+, ruAA | AA+.ru, AA.ru | AA+|ru|, AA|ru| |
| 2 | AA-(RU), A+(RU) | ruAA-, ruA+ | AA-.ru, A+.ru | AA-|ru|, A+|ru| |
| 3 | А (RU), A-(RU) | ruA, ruA- | A.ru, A-.ru | A|ru|, A-|ru| |
| 4 | Средний уровень кредитоспособности / финансовой надежности / финансовой устойчивости (от умеренно низкого до среднего) | BBB+(RU), BBB (RU) | ruBBB+, ruBBB | BBB+.ru, BBB.ru | BBB+|ru|, BBB|ru| |
| 5 | BBB-(RU), BB+(RU) | ruBBB-, ruBB+ | BBB-.ru, BB+.ru | BBB-|ru|, BB+|ru| |
| 6 | Низкий уровень кредитоспособности / финансовой надежности / финансовой устойчивости (от преддефолтного до умеренно низкого) | BB (RU) | ruBB | BB.ru | BB|ru| |
| 7 | BB-(RU) и ниже | ruBB- и ниже | BB-.ru и ниже | BB-|ru| и ниже |

*Источник: Банк России*

*Приложение 2*

*Таблица 2. Расчет используемых объясняющих переменных.*

|  |  |  |
| --- | --- | --- |
| Краткое наименование | Формула расчета/Определение | Группа факторов |
| UNEMP | Уровень безработицы | Макроэкономические |
| GDP | Темп роста ВВП (по отношению к соответствующему кварталу предыдущего года) |
| INFL | ИПЦ (по отношению к соответствующему кварталу предыдущего года) |
| DATE | - | Временной эффект |
| LI | 1, есть значимая исковая нагрузка; 0, иначе | Нефинансовые |
| MSP | 1, если относится к МСП; 0, иначе |
| SECTOR | - |
| CA\_CL | Оборотные Активы / Краткосрочные Обязательства | Ликвидность (финансовые) |
| C\_A | Денежные средства/Активы |
| NWC\_A | (Оборотные Активы - Краткосрочные Обязательства)/Активы |
| D\_EBIT | Долг/Операционная прибыль | Долговая нагрузка (финансовые) |
| C\_D | Денежные средства/Обязательства |
| D\_REV | Долг/Выручка |
| D\_E | Долг/Капитал |
| EBIT\_I | Операционная прибыль/Процентные расходы |
| ROA | Чистая прибыль/Активы | Рентабельность (финансовые) |
| ROE | Чистая прибыль/Капитал |
| NWC\_R | (Оборотные Активы - Краткосрочные Обязательства)/Выручка | Прочее (финансовые) |
| EBIT\_REV | Операционная прибыль/Выручка |
| EQ\_A | Капитал/Активы |
| RE\_ A | Валовая прибыль/Активы |
| EBIT\_A | Операционная прибыль/Активы |

*Источник: составлена автором*

*Приложение 2*

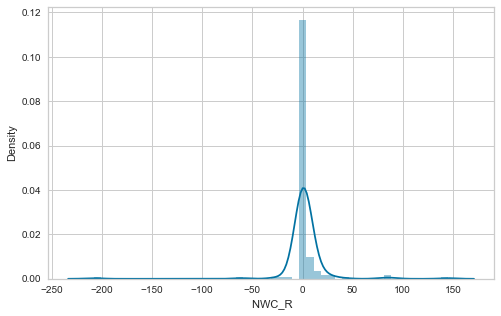
**

Рис.1 График распределения NWC\_R до обработки выбросов

Источник: расчеты автора

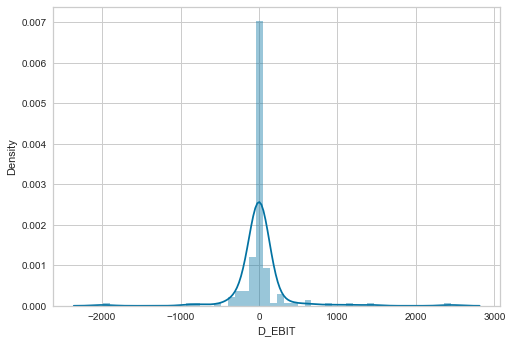
**

Рис.2 График распределения D\_EBIT до обработки выбросов

Источник: расчеты автора

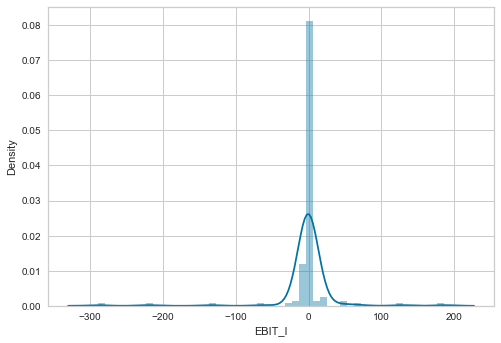
**

Рис.3 График распределения EBIT\_I до обработки выбросов

Источник: расчеты автора

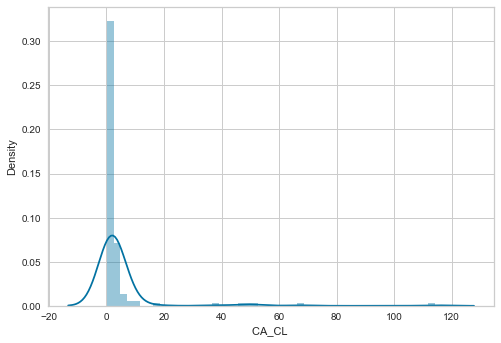
**

Рис.4 График распределения CA\_CL до обработки выбросов

Источник: расчеты автора

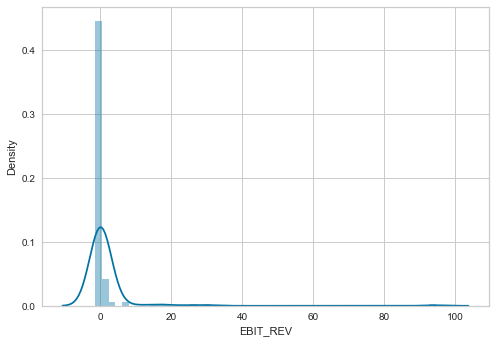
**

Рис.5 График распределения EBIT\_REV до обработки выбросов

Источник: расчеты автора

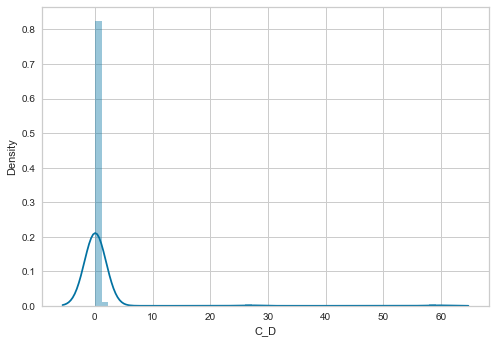
**

Рис.6 График распределения C\_D до обработки выбросов

Источник: расчеты автора

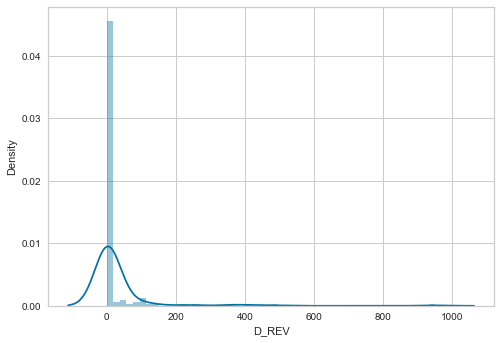
**

Рис.7 График распределения D\_REV до обработки выбросов

Источник: расчеты автора

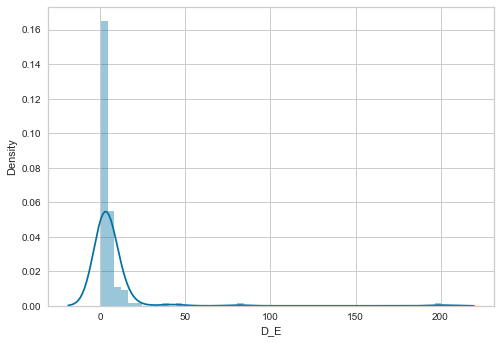
**

Рис.8 График распределения D\_E до обработки выбросов

Источник: расчеты автора

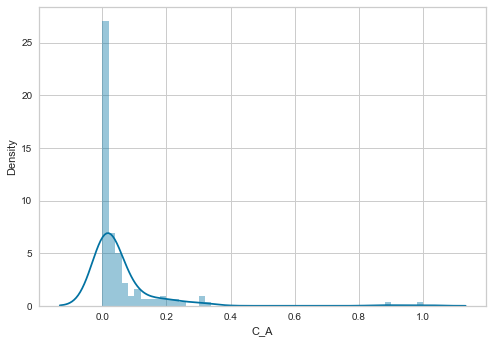
**

Рис.9 График распределения C\_A до обработки выбросов

Источник: расчеты автора

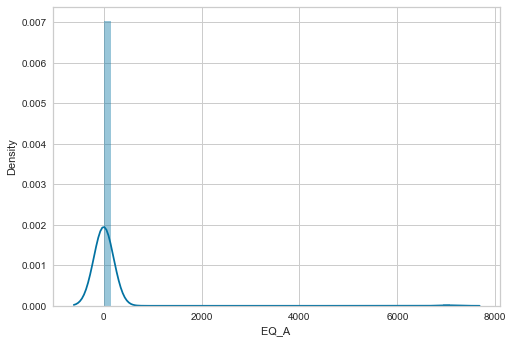
**

Рис.10 График распределения EQ\_A до обработки выбросов

Источник: расчеты автора

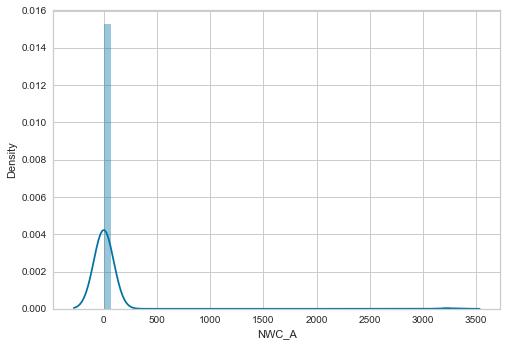
**

Рис.11 График распределения NWC\_A до обработки выбросов

Источник: расчеты автора

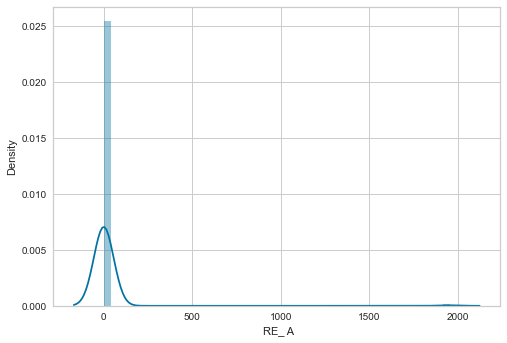
**

Рис.12 График распределения RE\_F до обработки выбросов

Источник: расчеты автора

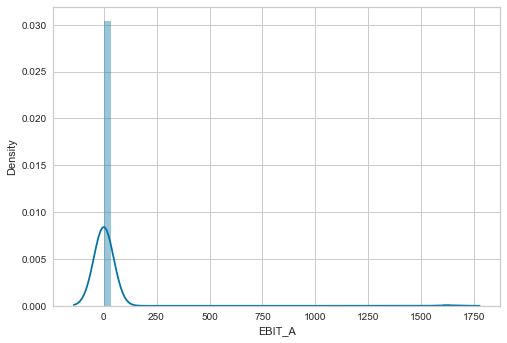
**

Рис.13 График распределения EBIT\_A до обработки выбросов

Источник: расчеты автора

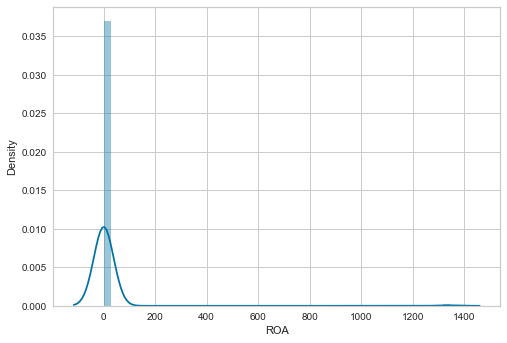
**

Рис.14 График распределения ROA до обработки выбросов

Источник: расчеты автора

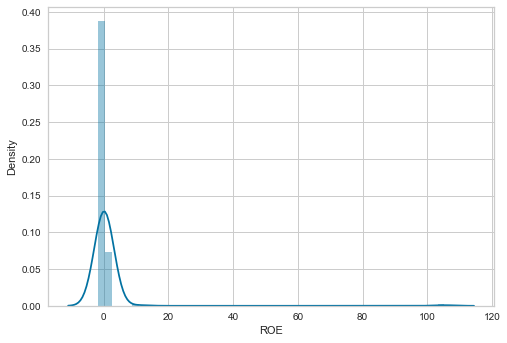
**

Рис.15 График распределения ROE до обработки выбросов

Источник: расчеты автора

*Приложение 4*

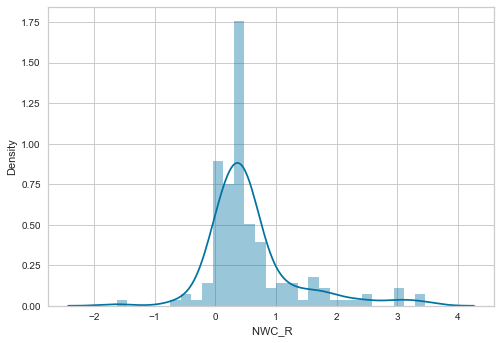
**

Рис.16 График распределения NWC\_R после обработки выбросов

Источник: расчеты автора

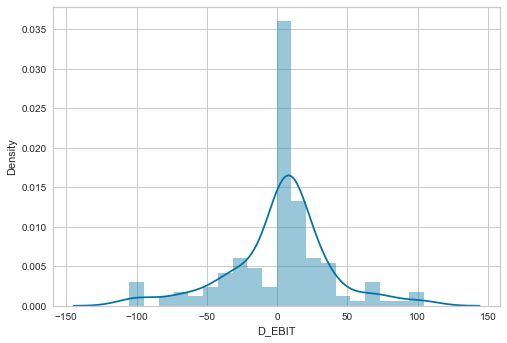
**

Рис.17 График распределения D\_EBIT после обработки выбросов

Источник: расчеты автора

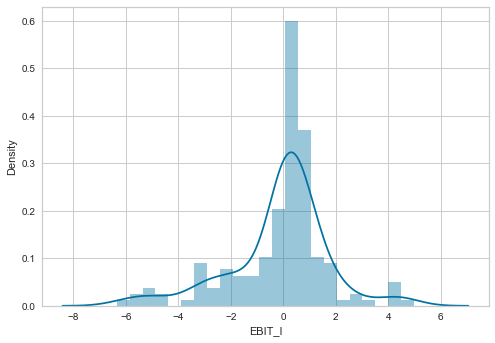
**

Рис.18 График распределения EBIT\_I после обработки выбросов

Источник: расчеты автора

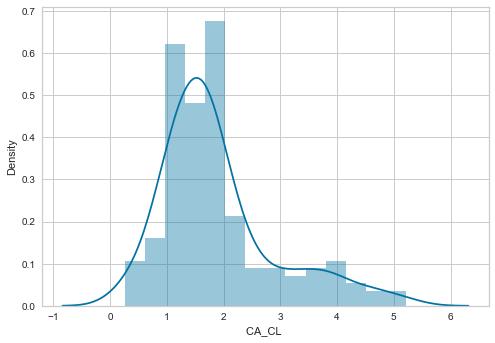
**

Рис.19 График распределения CA\_CL после обработки выбросов

Источник: расчеты автора

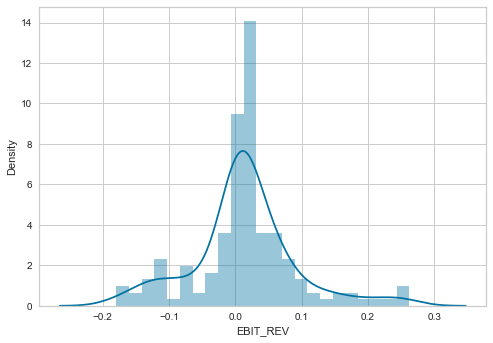
**

Рис.20 График распределения EBIT\_REV после обработки выбросов

Источник: расчеты автора

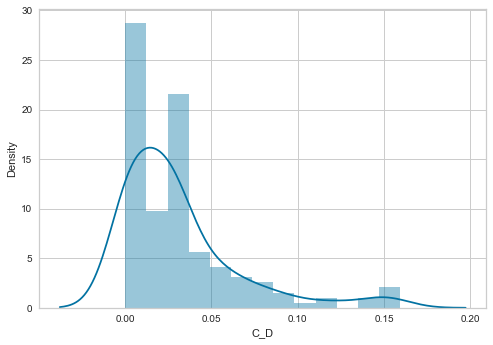
**

Рис.21 График распределения C\_D после обработки выбросов

Источник: расчеты автора

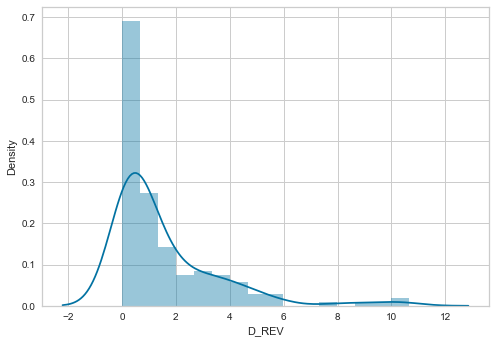
**

Рис.22 График распределения D\_REV после обработки выбросов

Источник: расчеты автора

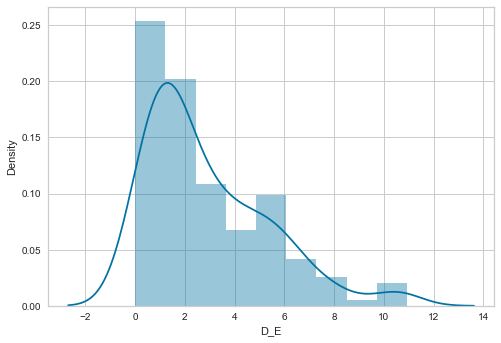
**

Рис.23 График распределения D\_E после обработки выбросов

Источник: расчеты автора

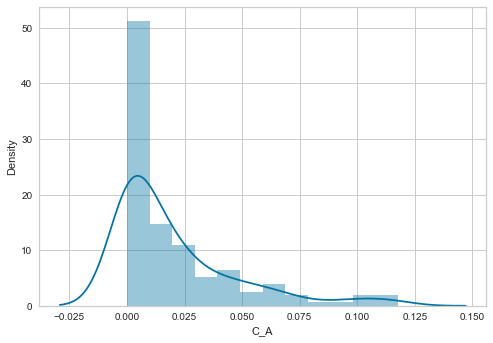
**

Рис.24 График распределения C\_A после обработки выбросов

Источник: расчеты автора

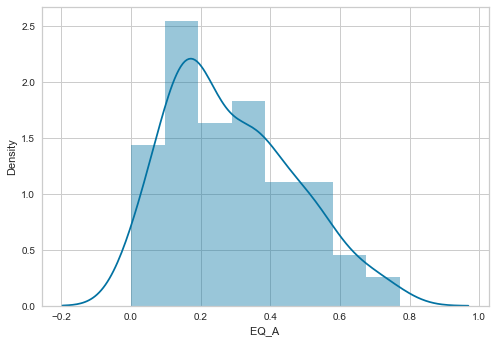
**

Рис.25 График распределения EQ\_A после обработки выбросов

Источник: расчеты автора

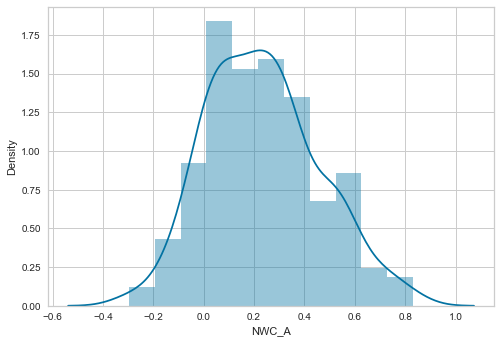
**

Рис.26 График распределения NWC\_A после обработки выбросов

Источник: расчеты автора

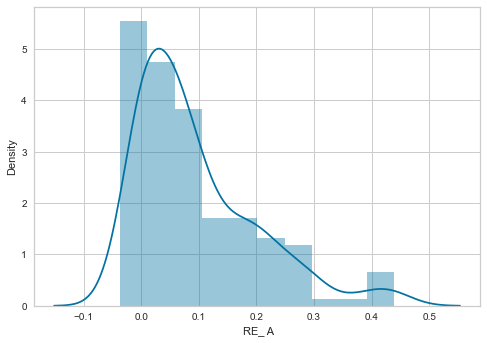
**

Рис.27 График распределения RE\_A после обработки выбросов

Источник: расчеты автора

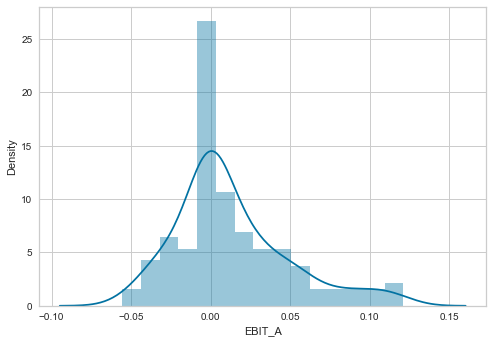
**

Рис.28 График распределения EBIT\_A после обработки выбросов

Источник: расчеты автора

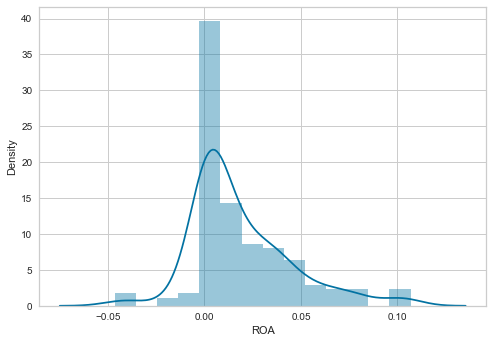
**

Рис.29 График распределения ROA после обработки выбросов

Источник: расчеты автора

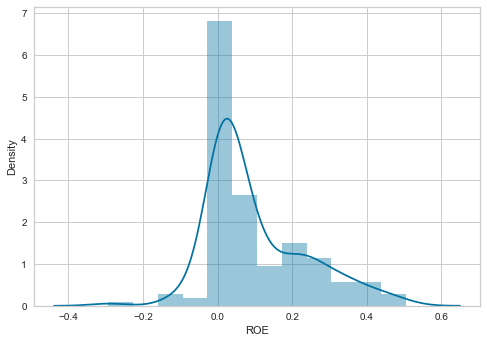
**

Рис.30 График распределения ROE после обработки выбросов

Источник: расчеты автора

*Приложение 5*

*Таблица 3. Корреляционная матрица используемых переменных*

**

*Источник: расчеты автора*

*Приложение 6*

*Таблица 4. Корреляционная матрица кредитного рейтинга и принадлежности к определенной отрасли*

|  |  |
| --- | --- |
|  | **CR** |
| **SECTOR\_Производство строительных материалов** | -0,16 |
| **SECTOR\_Строительство зданий** | -0,30 |
| **SECTOR\_Вспомогательная транспортная деятельность** | -0,13 |
| **SECTOR\_Телекоммуникационное оборудование** | -0,09 |
| **SECTOR\_FMCG ретейл** | -0,14 |
| **SECTOR\_Компьютеры, периферия и комплектующие** | -0,09 |
| **SECTOR\_Оптовая торговля** | 0,28 |
| **SECTOR\_Специализированная розничная торговля** | -0,15 |
| **SECTOR\_Оборудование и услуги для добычи нефти и газа** | -0,14 |
| **SECTOR\_Производство лекарств и биотехнологии** | -0,10 |
| **SECTOR\_Другие отрасли** | -0,14 |
| **SECTOR\_ИТ сервисы и программное обеспечение** | -0,08 |
| **SECTOR\_АПК и сельское хозяйство** | 0,17 |
| **SECTOR\_Связь и телекоммуникация** | 0,02 |
| **SECTOR\_Прочее машиностроение и приборостроение** | 0,17 |
| **SECTOR\_Полупроводники** | -0,04 |
| **SECTOR\_Образование** | 0,04 |
| **SECTOR\_Производство летательных аппаратов и вооружения** | -0,05 |
| **SECTOR\_Недвижимость и фонды недвижимости** | 0,06 |
| **SECTOR\_Утилизация и удаление отходов** | 0,06 |
| **SECTOR\_Строительство сооружений** | 0,05 |
| **SECTOR\_Цветная металлургия** | 0,10 |
| **SECTOR\_Производство продуктов и напитков** | 0,23 |
| **SECTOR\_Товары длительного пользования** | 0,15 |
| **SECTOR\_Общественное питание** | 0,10 |
| **SECTOR\_Тара и упаковка** | 0,10 |
| **SECTOR\_Персональные и коммерческие услуги** | 0,10 |

*Источник: расчеты автора*